DOCUMENT RESUME

ED 211 427 SO 013 806

TITLE	Procedural Handbook: 1978-79 Art Assessment.
INSTITUTION	Education Commission of the States, Denver, Colo. National Assessment of Educational Progress.
SPONS AGENCY	National Inst. of Education (ED), Washington, D.C.
REPORT NO	ISBN-0-89398-014-5; NAEP-10-A-40
PUB DATE	81
GRANT	NIE-G-80-0003
NOTE	86p.; For a related document, see ED 186 331.
AVAILABLE FROM	National Assessment of Educational Progress, 1860 Lincoln St., Suite 700, Denver, CO 80295 ($8.60).
EDRS PRICE	MF01/PC04 Plus Postage.
DESCRIPTORS	*Administration; Art Appreciation; *Art Education; Art History; Cognitive Ability; Data Analysis; Data Collection; *Educational Assessment; Educational Objectives; Elementary Secondary Education; *Evaluation Methods; Multiple Choice Tests; Sampling; Scoring; Student Attitudes
IDENTIFIERS	*National Assessment of Educational Progress; Second Art Assessment (1979)

ABSTRACT

This handbook describes the procedures used to develop, administer, and analyze the results of the 1978-79 art assessment of 9-year-olds, and 17-year-olds by the National Assessment of Educational Progress (NAEP). The primary purpose of the handbook is to provide detailed procedural information for people interested in replicating the assessment or in need of more information than is provided in the reports containing assessment data. The seven chapters cover objectives redevelopment, exercise creation, preparation of assessment booklets, sampling data collection, scoring, and data analysis. Each chapter explains the basic procedures used for the 1978-79 art assessment and contrasts the procedures to those used in earlier years (if there were changes). Appendices include definitions of reporting groups used by NAEP, forms used to gather background information about students and schools, response rates, computation of achievement measures, and procedures for smoothing respondent weights. A glossary of National Assessment terms is provided at the end of the book. Primary type of information provided by report: Procedures (Overview). (Author/RM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED211427

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☒ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

PROCEDURAL HANDBOOK:

1978-79 ART ASSESSMENT

Report No. 10-A-40

by the
National Assessment of Educational Progress

Education Commission of the States
Suite 700, 1860 Lincoln Street
Denver, Colorado 80295

December 1981

SP013806

The National Assessment of Educational Progress is funded by the National Institute of Education under a grant to the Education Commission of the States. It is the policy of the Education Commission of the States to take affirmative action to prevent discrimination in its policies, programs and employment practices.

Library of Congress
Catalog Card Number: 72-183665

Although a few early National Assessment reports have individual catalog card numbers, all recent reports have been assigned the above series number.

ISBN 0-89398-014-5

The National Assessment of Educational Progress is an education research project mandated by Congress to collect and report data, over time, on the performance of young Americans in various learning areas. National Assessment makes available information on assessment procedures and materials to state and local education agencies and others.

The work upon which this publication is based was performed pursuant to Contract No. OEC-0-74-0506 of the National Center for Education Statistics and the National Institute of Education; also, Grant No. NIE-G-80-0003 of the National Institute of Education. It does not, however, necessarily reflect the views of those agencies.

TABLE OF CONTENTS

LIST OF TABLES..... v

FOREWORD..... vii

ACKNOWLEDGMENTS..... ix

INTRODUCTION..... xi

CHAPTER 1 Objectives Redevelopment..... 1
Outline of Art Objectives..... 3

CHAPTER 2 Development of Exercises..... 5
Exercise Development..... 5
Field Tryout Procedures..... 6

CHAPTER 3 Preparation of Assessment Materials..... 7
Preparation of Booklets and Audio Tapes..... 7
Differences in Item Booklets in the 1974-75 and 1978-79 Assessments..... 8

CHAPTER 4 Sampling..... 11
Overview of the National Assessment Sample Design..... 12
Survey Weights..... 12

CHAPTER 5 Data Collection..... 15

CHAPTER 6 Scoring..... 19

CHAPTER 7 Data Analysis..... 23
Measures of Achievement..... 23
Estimating Variability in Achievement Measures..... 25
Controlling Nonrandom Errors..... 26

APPENDIX A Definitions of National Assessment Reporting Groups..... 27
Group Definitions..... 27

APPENDIX B Forms Used to Obtain Background Information..... 31

APPENDIX C Response Rates for Assessment Samples..... 55

APPENDIX D Computation of Measures of Achievement, Changes in Achievement and Standard Errors..... 59
Measures of Achievement..... 59
Computation of Standard Errors..... 61

APPENDIX E Adjustment of Respondent Weights by Smoothing to Reduce Random Variability of Estimated Population Proportions..... 65
Background..... 65
Smoothing Procedures Used by National Assessment..... 66
The Current Smoothing Procedure..... 67
Adjustment of Weights by Users..... 68
Changes in Smoothed Proportions as New Assessments Are Completed..... 70

GLOSSARY OF NATIONAL ASSESSMENT TERMS..... 71

BIBLIOGRAPHY..... 77

LIST OF TABLES

TABLE 1. Desired Percent of Assessment Exercise Time by Objective and Subobjectve, for Ages 9, 13 and 17, 1978-79 Assessment..... 2

TABLE 2. Desired Percent of Assessment Exercise Time by Age and Major Objective, 1974-75 Assessment..... 2

TABLE 3. Number of PSUs and Schools Within PSUs Selected in 1974-75 and 1978-79..... 13

TABLE 4. School Cooperation Rates, 1978-79 Assessment..... 15

TABLE 5. Number of Students Responding to Each Item Booklet in 1978-79 Assessment, by Age..... 16

TABLE 6. Between-Scorer Quality-Control Summaries for Art Open-Ended Scoring Summarized Across Seven Weeks..... 21

TABLE 7. Within-Scorer Quality-Control Summaries for Art Open-Ended Scoring..... 21

TABLE C-1. Number of Students Assessed and Percent of Sample Covered, by Age and Assessment Year..... 55

TABLE D-1. Average Number of Respondents in Reporting Groups Taking an Item Booklet, by Age and Assessment Year..... 63

TABLE D-2. Estimated Current Population Proportions of National Assessment Reporting Groups for In-School Students..... 64

TABLE E-1. Smoothing Cells Used for the 1978-79 Smoothing Procedure..... 67

TABLE E-2. Smoothed Frequencies From 10-Year Smooth by Smoothing Cell and Year for 9-Year-Olds..... 68

TABLE E-3. Smoothed Frequencies From 10-Year Smooth by Smoothing Cell and Year for 13-Year-Olds..... 69

TABLE E-4. Smoothed Frequencies From 10-Year Smooth by Smoothing Cell and Year for In-School 17-Year-Olds..... 69

FOREWORD

When the U.S. Office of Education was chartered in 1867, one charge to its commissioners was to determine the nation's progress in education. The National Assessment of Educational Progress (NAEP) was initiated a century later to address, in a systematic way, that charge.

Since 1969, the National Assessment has gathered information about levels of educational achievement across the country and reported its findings to the nation. It has surveyed the attainments of 9-year-olds, 13-year-olds, 17-year-olds and sometimes adults in art, career and occupational development, citizenship, literature, mathematics, music, reading, science, social studies and writing. All areas have been periodically reassessed in order to detect any important changes. Including students who participated in the latest art assessment, National Assessment has interviewed and tested more than 1,000,000 young Americans.

Learning-area assessments evolve from a consensus process. Each assessment is the product of several years of work by a great many educators, scholars and lay persons from all over the nation. Initially, these people design objectives for each subject area, proposing general goals they feel Americans should be achieving in the course of their education. After careful reviews, these objectives are given to exercise (item) writers, whose task it is to create measurement instruments appropriate to the objectives.

When the exercises have passed extensive reviews by subject-matter specialists, measurement experts and lay persons, they are administered to probability samples. The people who compose these samples are chosen in such a way that the results of their assessment can be generalized to an entire national population. That is, on the basis of the performance of about 2,500 9-year-olds on a given exercise, we can make generalizations about the probable performance of all 9-year-olds in the nation.

After assessment data have been collected, scored and analyzed, the National Assessment publishes reports and disseminates the results as widely as possible. Not all exercises are released for publication. Because NAEP will readminister some of the same exercises in the future to determine whether the performance levels of Americans have increased, remained stable or

decreased, it is essential that they not be released in order to preserve the integrity of the study.

ACKNOWLEDGMENTS

Many organizations and individuals have made substantial contributions to the art assessments. Not the least of those to be gratefully acknowledged are the administrators, teachers and students who cooperated so generously during the collection of the data.

Special acknowledgment must go to the many art educators and specialists who provided their expertise in the development, review and selection of the assessment objectives and exercises. Development of the art assessment was coordinated by Sarah Knight.

Administration of the art assessment was conducted by the Research Triangle Institute, Raleigh, North Carolina. Scoring and processing were carried out by the Measurement Research Center (now Westinghouse DataScore Systems), Iowa City, Iowa, and by the National Assessment staff.

The actual preparation of this report was a collaborative effort of the National Assessment staff. Special thanks must go to Donald T. Searls for information on sampling and data analysis and to Marci Reser and Deborah Houy for production. Scoring and statistical analyses for the art assessment were supervised by Sarah Knight and Donald Phillips.

Roy H. Forbes
Director

INTRODUCTION

The National Assessment of Educational Progress (NAEP) has completed two assessments of art, the first conducted in 1974-75 and the second in 1978-79. Each assessment surveyed the achievement and attitudes of American 9-, 13- and 17-year-olds, using a deeply stratified, multistage probability sample design. This report documents procedures used in the 1978-79 art assessment and also describes changes in procedures between the assessments.

To measure changes in performance between 1974-75 and 1978-79, approximately half of the exercises assessed in the first assessment were reassessed in the second under almost identical administrative conditions. To measure the status of art achievement in 1978-79, National Assessment consultants reviewed the objectives used in the first assessment and developed additional exercises to provide wider coverage of those objectives.[1]

Approximately 7,500 9-year-olds, 11,000 13-year-olds and 13,500 17-year-olds participated in the 1978-79 art assessment. Because there were more art exercises for 13-year-olds than available assessment space, six exercises were held and administered in the next year's assessment (1979-80). During the 1979-80 assessment, 2,749 13-year-olds responded to these six art exercises.

Since National Assessment reports results for groups of students, not individuals, it is not necessary for each student to respond to every item (exercise).[2] Each respondent completed only one item booklet of about 45 minutes in length. Between 2,400 and 2,800 students responded to each booklet. In 1978-79, three exercise booklets for 9-year-olds, four booklets for 13-year-olds and five booklets

---

[1]Although National Assessment traditionally revises objectives prior to an assessment, a shortage of funds and time prohibited complete redevelopment of the art objectives. Previous art objectives were reviewed by consultants and found to be usable with some shifts in emphasis.

[2]National Assessment uses the term "exercise" to mean an assessment item. The terms "exercise" and "item" are used interchangeably in this report.

for 17-year-olds contained art exercises. In the 1979-80 assessment, only one item booklet for 13-year-olds included art items.

In both the 1974-75 and the 1978-79 assessments, 13-year-olds were assessed in October through December, 9-year-olds in January and February and 17-year-olds in March and April. Thus, the amount of school experience for each age group was approximately the same in each assessment.

The exercises for each assessment were administered by a professional data collection staff to minimize the burden on participating schools and to maximize the uniformity of assessment conditions. Instructions and items were recorded on a paced audio tape and played back to students to reduce the potential effect of reading difficulties and to ensure that all students moved through the packages at the same speed.

The majority of the art items were multiple-choice; a few exercises from the first assessment that were reassessed in 1978-79 were open-ended. Many items included more than one part.

Multiple-choice items were scored by an optical scanning machine; open-ended items were hand-scored by trained scorers using scoring guides developed to define categories of acceptable and unacceptable responses.

National Assessment reports estimated percentages of correct responses for single items. When a report indicates that "85% of the 13-year-olds gave a correct response," it means that an estimated 85% of the 13-year-olds would have given a correct response if all the 13-year-olds in schools across the country had been assessed. National Assessment also aggregates percentages of success on various sets of items to provide data on changes in performance between assessments and on the differential performance of population subgroups. In addition to reporting national results, National Assessment provides data on the performance of various population subgroups within the national population, defined by sex, race/ethnicity, region of the country, size and type of community lived in and level of parental education. In addition, for the art assessments, some data are available on the extent of art training and/or experience with art -- either formal or informal -- that 9-, 13- and 17-year-olds have had.

This handbook describes the procedures used to develop, administer and analyze the results of the 1978-79 art assessment. The primary purpose of this handbook is to provide detailed procedural information for people interested in replicating the assessment or in need of more information than is provided in the reports containing assessment data. The seven chapters cover objectives redevelopment, exercise creation,

preparation of assessment booklets, sampling, data collection, scoring and data analysis. Each chapter explains the basic procedures used for the 1978-79 art assessment and contrasts the procedures to those used in earlier years (if there were changes).

Appendixes include definitions of reporting groups used by National Assessment, forms used to gather background information about students and schools, response rates, computation of achievement measures and procedures for smoothing respondent weights. A glossary of National Assessment terms is provided at the end of the book.

The objectives for the art assessments, Art Objectives (1971) and a one page objectives supplement and released exercises, The Second Assessment of Art, 1978-79: Released Exercise Set (1980), are available from NAEP.

A supplement to the released exercise set that will include released open-ended exercises, their scoring guides and some selected data is planned for publication in late 1981. By the end of 1981, National Assessment also will complete reports on the results of the 1978-79 art assessment and the changes in art performance between the 1974-75 and 1978-79 assessments. For the exact titles of these publications and information about their availability, more complete data or other assistance concerning the art assessments, please contact National Assessment.

CHAPTER 1

OBJECTIVES REDEVELOPMENT

The primary goals of the National Assessment of Educational Progress (NAEP) are to report on the current education status of young Americans and to monitor any changes in achievement over time. For each learning area to be assessed, NAEP asks consultants to develop objectives that define the subject area. Since the objectives provide guidelines for exercise developers, consultants are asked to include examples of the knowledge, skills and attitudes to be assessed at each age level.

Education in America is a collaborative enterprise involving a great many people with widely differing philosophies. Providing information about education nationwide would be considerably easier if there were a consensus about the means and ends of American education, but the fact is that Americans have conflicting and sometimes contradictory values regarding the goals of education and the means for achieving them. To develop an assessment that is truly national in scope and takes into account the diversity of curricula, values and goals across the country, National Assessment employs a consensus process for developing objectives, with representation of many different groups of people.

Several types of consultants helped develop the objectives for the first assessment of art. College and university specialists in art, classroom teachers, curriculum supervisors and persons involved in teacher education made sure that the objectives included important concepts, skills and attitudes that the schools currently were teaching and that they should be teaching. Concerned citizens, parents and other interested lay persons also had to agree that the objectives were important for young people to achieve, were free of education jargon and were not biased against or offensive to any groups. Consultants were selected to represent different regions of the country and minority groups. They also represented a range of experience with students of different ages and community types.

In preparation for the 1978-79 art assessment, the 1974-75 art objectives were reviewed by art educators. A shortage of funds and a very tight schedule for completing the assessment made a complete redevelopment of the objectives impractical. To reflect

shifts of emphasis in art education, the objectives were reviewed and the amount of assessment time to be devoted to each objective was modified as necessary. Suggestions for more extensive revisions were provided to National Assessment for future implementation.

As seen in Tables 1 and 2, for the 1978-79 art assessment, the objectives were assigned the same percentage of time, or weight, for each age, although the weights used in 1974-75 varied with age for some objectives. The major shift between assessments was that the production objective (III) received more emphasis for 13- and 17-year-olds, with the amount of time for objectives concerning knowledge about art (IV) and aesthetic judgments (V) correspondingly reduced.

Although in 1974-75 percentages of time were assigned only to major objectives, in 1978-79 percentages were assigned to subobjectives as well, to give additional guidance to exercise developers.

An outline of the art objectives used in 1974-75 and 1978-79 follows. For addi-

TABLE 1. Desired Percent of Assessment Exercise Time by Objective and Subobjective, for Ages 9, 13 and 17, 1978-79 Assessment

	Objective				
Subobjective	I	II	III	IV	V
A	8.0%	8.0%	6.0%	3.0%	3.0%
B	12.0	4.8	6.0	3.0	3.0
C	NA+	2.4	6.0	3.0	3.0
D	NA	2.4	6.0	3.0	6.0
E	NA	2.4	6.0	3.0	NA
Total	20.0	20.0	30.0	15.0	15.0

+Not assessed.

TABLE 2. Desired Percent of Assessment Exercise Time by Age and Major Objective, 1974-75 Assessment

	Objective				
	I	II	III	IV	V
Age 9	20.0%	20.0%	30.0%	15.0%	15.0%
Age 13	20.0	20.0	25.0	17.5	17.5
Age 17	20.0	20.0	20.0	20.0	20.0

tional detail concerning art objectives, see Art Objectives (1971).

Outline of Art Objectives

I. PERCEIVE AND RESPOND TO ASPECTS OF ART

Aspects of art are defined as: sensory qualities of color, line, shape and texture; compositional elements such as structure, space, design, balance, movement, placement, closure, contrast and pattern; expressive qualities such as mood, feeling and emotion; subject matter, including (1) objects, themes (the general subject of a work, i.e., landscape or battle scene), events and ideas (general presymbolic meanings) and (2) symbols and allegories; and expressive content, which is a unique fusion of the foregoing aspects

A. Recognize and describe the subject-matter elements of works of art

B. Go beyond the recognition of subject matter to the perception and description of formal qualities and expressive content (the combined effect of the subject matter and the specific visual form that characterizes a particular work of art)

II. VALUE ART AS AN IMPORTANT REALM OF HUMAN EXPERIENCE

A. Be affectively oriented toward art

B. Participate in activities related to art

C. Express reasonably sophisticated conceptions about and positive attitudes toward art and artists

D. Demonstrate an open-mindedness toward different forms and styles of art

E. Demonstrate an open-mindedness toward artistic experimentation

III. PRODUCE WORKS OF ART

A. Produce original and imaginative works of art

B. Express visual ideas fluently

C. Produce works of art with a particular composition, subject matter, expressive character or expressive content

D. Produce works of art that contain various visual conceptions

E. Demonstrate knowledge and application of media, tools, techniques and forming processes

IV. KNOW ABOUT ART

A. Recognize major figures and works in the history of art and understand their significance. (Significance as it is used here refers to such things as works of art that began new styles, markedly influenced subsequent works, changed the direction of art, contained visual and technical discoveries, expressed particularly well the spirit of their age and those considered to be the major works of major artists)

B. Recognize the styles of art, understand the concept of style and analyze works of art on the basis of style

C. Know the history of art activity and understand the relation of one style or period to other styles and periods

D. Distinguish between factors of a work of art that relate principally to the personal style of the artist and factors that relate to the stylistic period or the entire age

E. Know and recognize the relationships that existed between art and the other disciplines of the humanities (literature, music and particularly the history of ideas and philosophy) during a given period

V. MAKE AND JUSTIFY JUDGMENTS ABOUT THE AESTHETIC MERIT AND QUALITY OF WORKS OF ART

Statements of aesthetic quality are those that characterize the various aspects of a work of art, while statements of aesthetic merit are assertions about the degree of goodness or badness of the work. Justifications of aesthetic merit are based on criteria such as the degree to which the work is integrated and whether contact with the work results in a vivid and fused experience

A. Make and justify judgments about aesthetic merit

B. Make and justify judgments about aesthetic quality

C. Apply specific criteria in judging works of art

D. Know and understand criteria for making aesthetic judgments

CHAPTER 2

DEVELOPMENT OF EXERCISES

Exercise Development

Exercise development for the 1978-79 art assessment began with a planning conference early in 1977. At this conference, consultants evaluated exercises developed for the first assessment but not assessed and determined which objectives were not adequately measured by existing materials. In addition, prototype items that could be used as models by exercise writers were identified.

Consultants who are experts in art, child development and art education wrote new exercises during March 1977. The new exercises and those chosen from the items developed for the first assessment but not used were field tested in spring 1977. The exercises and the results of the field trials were reviewed by consultants during summer 1977. This review provided guidance for further exercise development, which was conducted in late summer and early fall 1977. Exercises from this development were field tested during October and November. Following these field tests, all newly developed exercises and field test results were reviewed by subject-matter experts and lay persons.

Subject-matter specialists reviewed exercises to ensure that they were appropriate for the age level being assessed and that the content included was correct. Lay citizens, representing a variety of occupations and interests, also reviewed the exercises, checking for sex or racial/ethnic bias and considering the general importance of each exercise.

By the end of February 1978, exercises had been revised and edited and were in final form for inclusion in the assessment. Since more items were developed than could be assessed, in March the items to be used for the first time in the 1978-79 assessment were selected. Exercises to be used were selected at a conference of subject-matter experts and educators. Items from the first assessment to be used to measure changes in performance were automatically included in the second assessment. Using the weighted objectives as a guide to the amount of assessment time to be used for each subobjective, exercises (including those from the first assessment) were selected to approximate these amounts.

Field Tryout Procedures

The field trials for the 1978-79 art assessment were performed in schools across the country to discover potential problems with items in wording, directions or administrative procedures and to collect item statistics, timing information and scoring information. "Tryout" schools were selected to represent high- and low-income communities as well as more typical communities. The items being field tested were administered to students in at least four classrooms (approximately 100 students) at each age for which an item was intended. In order to simulate actual assessment field procedures, students recorded their answers in the test booklets; directions and questions were read to students from an audio tape; and National Assessment staff members, rather than classroom teachers, administered the test. The completed "tryout" packages were then scored and results were analyzed. The tryout data, as well as the administrators' reports of any field problems, helped both NAEP staff and consultants to evaluate and revise the exercises.

CHAPTER 3

PREPARATION OF ASSESSMENT MATERIALS

Preparation of Booklets and Audio Tapes

National Assessment uses a matrix sampling approach, with different nationally representative samples of students responding to different item booklets (see Chapter 4 for details). Since the Assessment's aim is to describe results for groups of students (males, blacks, students in the West, and so on), not individuals, it is not necessary for each student to respond to all the items. Each student responded to one booklet of items designed to be completed in a single class period.

The 1978-79 assessment was a combined assessment of art, music and writing. Because of the length of many of the art and writing exercises, only two subject areas were included in an item booklet. Booklets of exercises included either music and writing exercises, art and writing exercises, or art exercises only.

Following the selection of art exercises, National Assessment staff grouped and sequenced them into exercise booklets. Since students at different ages received somewhat different sets of exercises, booklets were constructed separately for each age level. Thus, exercises for 9-year-olds were not sequenced in the same order as those for 13-year-olds, and so forth.

In 1978-79, there were three exercise booklets that contained art exercises for 9-year-olds, four such booklets for 13-year-olds and five such booklets for 17-year-olds. One booklet for 13-year-olds in the 1979-80 assessment included art items.

The following constraints were observed in preparing exercise booklets:

1. Each booklet contained exercises of varying difficulty so that students would not become bored by many easy exercises or discouraged by many difficult exercises.

2. Exercises could not cue other exercises. In other words, the answer to one exercise could not be contained in another exercise in the same booklet.

3. Each booklet was timed so that it would take no more than 45 minutes of a student's time -- the length of a typical class period. Booklets contained approx-

imately 30-35 minutes of exercise time and an additional 10-15 minutes of introductory material, instructions and background questions.

4. Booklets were designed to be, insofar as possible, parallel with respect to the number of different objectives measured and difficulty levels. Items measuring a particular objective were scattered throughout the booklets so that many different students would respond to questions related to a particular objective.

5. At ages 13 and 17, booklets containing reassessed items used to measure change between 1974-75 and 1978-79 were identical in content and item sequence in both assessments. The six exercises for 13-year-olds assessed in 1979-80 (the year following the main art assessment) were in the same booklet. None of these items was used to measure change in achievement from 1974-75.

National Assessment makes every effort to minimize difficulties connected with the testing situation so that results will be, as nearly as possible, an accurate reflection of what students know and can do. For example, students marked their answers directly in the assessment booklets, not on separate answer sheets. It was felt that this procedure would reduce possibilities for errors in marking answer sheets, especially for younger students. To minimize guessing, students were encouraged to select the "I don't know" response option included with some multiple-choice items or to write "I don't know" on the answer line for open-ended questions if they felt they did not know the answer to a question. Many multiple-choice art exercises asked for a student's opinion about a work of art. Such items did not include the "I don't know" option.

Paced audio tapes were prepared for each exercise booklet. Instructions, most of the written portions of an exercise stimulus and response options were read aloud to minimize the effect of any reading difficulties and to ensure that all students moved through the booklets at the same speed. During the field testing of the exercises, administrators had determined the time needed for most students to respond to an item, and this amount of time was allowed on the audio tape. In addition, the use of tapes helped to ensure uniform assessment conditions across the country.

Differences in Item Booklets in the 1974-75 and 1978-79 Assessments

National Assessment makes every effort to make assessment conditions for items measuring change identical from assessment to assessment so that any changes observed

will be attributable to changes in achievement rather than a response to an altered testing condition.

As noted previously, booklets of items used to measure changes in achievement for 13- and 17-year-olds included identical items in the same sequence in both assessments. Items used for the first time in 1978-79 were included in other booklets. Tapescripts for the audio tapes accompanying the booklets used to measure change between assessments were also identical in each assessment. At age 9, booklets including items used to measure change also contained newly developed items; therefore, booklets were not the same for the two assessments. However, individual items used to measure change and tapescripts for those items were identical in both assessments. It was felt that the difference in item context probably would not have any appreciable effect on results. Items for 13-year-olds included in the 1979-80 assessment were newly developed for the second assessment and thus were not used to measure changes in performance.

CHAPTER 4

SAMPLING

This chapter gives an overview of the procedures used in designing and selecting the samples for the 1974-75 and 1978-79 art assessments. Sample design and selection for both assessments were conducted by the staff of the Research Triangle Institute, Raleigh, North Carolina, and monitored by National Assessment staff.

The target populations for the 1978-79 assessment consisted of 9-, 13- and 17-year-olds[1] enrolled in either public or private schools at the time of the assessment who were not functionally handicapped to the extent that they could not participate in an assessment. Other specific groups excluded were: non-English-speaking persons, those identified as nonreaders, persons physically and mentally unable to respond, and persons in institutions or attending schools established for the physically or mentally handicapped.

National Assessment did not follow up specific individuals from one assessment to the next. In other words, the students who participated in the 1974-75 assessment were not the same ones who participated in 1978-79. However, in each assessment year, participants were carefully selected to represent each age level. For example, although different sets of probability samples of 9-year-olds were used for the two assessments, each set contained nationally representative samples of students who were 9 years old during that assessment year. Thus, if we say that 9-year-olds' achievement declined between 1975 and 1979, we mean that students who were 9 years old in 1975 correctly answered the same questions more often than those who were 9 years old in 1979.

The definitions of the in-school target populations were identical in each assessment. The National Assessment samples were designed to provide approximately 2,500 respondents per exercise. These numbers allow reporting of data for the nation and for the

---

[1]Definitions of 1978-79 assessment age groups are: 9-year-olds -- born during the calendar year 1969; 13-year-olds -- born during the calendar year 1965; and 17-year-olds -- born October 1, 1961, through September 30, 1962.

subgroups defined in Appendix A.

There were minor differences in the sample designs for the 1974-75 and 1978-79 art assessments. A major difference in the samples obtained was that 17-year-olds who were dropouts or early graduates were assessed in 1974-75 but not in 1978-79, because of budget limitations. To ensure comparability, all comparisons of 17-year-olds' 1974-75 and 1978-79 results were made using data only for students attending school. All other differences are technical improvements or changes that increased sampling efficiency and lowered administrative costs, but did not affect comparability of the samples.

Overview of the National Assessment Sample Design

For all of its assessments, National Assessment uses a deeply stratified, three-stage national probability sample design with oversampling of low-income and rural areas. In the first stage, the United States is divided into geographical units of counties or groups of contiguous counties meeting a minimum population size requirement. These units, called primary sampling units (PSUs), are stratified by region and size of community. From the list of PSUs, a sample of PSUs is drawn without replacement with probability proportional to population size measures, representing all regions and sizes of communities. Oversampling of low-income and extreme-rural areas is first performed at this stage by adjusting the estimated population size measures of those areas to increase sampling rates. Within PSUs, Census Employment Survey Data are used to delineate and oversample low-income areas. Counties with high proportions of rural families are also oversampled.

In the second stage, all public and private schools within each PSU selected in the first stage are listed. Schools within each PSU are selected without replacement with probabilities proportional to the number of age-eligibles in the school.

The third stage of sampling occurs during the data collection period. A list of all age-eligible students within each selected school is made. A simple random selection of eligible students without replacement is obtained, and item booklets are administered to the selected students. Specially trained personnel select the sample and administer the booklets.

Survey Weights

The number of PSUs, schools within PSUs and students within schools is determined by optimum sampling principles. That is, a sample design is utilized that attempts to achieve the maximum precision for a given level of resources. Table 3 displays the number of PSUs used and the number of schools

TABLE 3. Number of PSUs and Schools Within PSUs Selected in 1974-75 and 1978-79

	1974-75 Assessment		1978-79 Assessment	
	No. of PSUs	No. of Schools	No. of PSUs	No. of Schools
Age 9	115	1,103	75	648
Age 13	115	972	75	650
Age 17	115	830	75	534

in which assessment sessions were conducted, by age, for the 1974-75 and 1978-79 assessments. Appendix C gives information about the number of students assessed.

Each respondent in the sample did not have the same probability of selection because some subpopulations were oversampled and because adjustments were made to compensate for some schools' refusals to participate and for student nonresponse. The selection probability for each individual was computed, and its reciprocal was used to weight each response in any statistical calculation to compensate for unequal rates of sampling and to ensure proper representation in the population structure. Procedures used to assign weights are discussed in Chapter 7 and Appendix E.

CHAPTER 5

DATA COLLECTION

National Assessment subcontracted data collection to the Research Triangle Institute, Raleigh, North Carolina. A professional data collection staff was used rather than school personnel to minimize the burden on participating schools and to ensure, insofar as possible, uniform administrative conditions across the country (Final Report...In-School Field Operations..., 1979).

Participation in the National Assessment is voluntary. NAEP makes every effort to encourage the schools selected in the sample to participate in the assessment, and National Assessment and Research Triangle Institute staffs have obtained high rates of school cooperation, as shown in Table 4 (Final Report...Field Operations..., 1979, p. 39, Table 27). Student cooperation rates were also high. The effect of student nonresponse is discussed in Appendix C. Table 5 shows the actual number of students that responded to a particular exercise booklet at each age level in the 1978-79 assessment.

TABLE 4. School Cooperation Rates, 1978-79 Assessment

Age	Percent of Eligible Schools Participating in 1978-79 Assessment
9	90.4
13	90.9
17	92.9
Overall	91.3

Each age group was assessed at approximately the same time of the school year in each assessment. As noted previously, 13-year-olds were assessed in October-December, 9-year-olds in January-February and 17-year-olds in March-May.

In 1978-79, booklets were administered to groups of 10-25 students, with each group responding to only one of the booklets for their age level. The groups varied in size depending on the number of eligible students and on an estimate of the rate of nonresponse for a particular school. In 1974-75, the planned session sizes were fixed at 12 students, with up to 4 alternates used if less than 12 students appeared for the session.

In each assessment, National Assessment takes steps to guarantee the anonymity of

TABLE 5. Number of Students Responding to Each Item Booklet in 1978-79 Assessment, by Age

Age 9		Age 13*		Age 17	
Booklet	Number Responding	Booklet	Number Responding	Booklet	Number Responding
1+	2,532	1+	2,755	1+	2,730
2+	2,553	2+	2,801	2+	2,746
3+	2,475	3+	2,775	3+	2,761
4+	2,494	4+	2,791	4+	2,772
5+	2,479	5+	2,785	5+	2,684
6+	2,522	6+	2,748	6+	2,739
7+	2,531	7+	2,736	7+	2,642
8+	2,524	8+	2,779	8+	2,656
9	2,486	9+	2,754	9++	2,787
10	2,483	10	2,758	10+	2,697
11	2,526	11	2,751	11	2,628
		12	2,720	12	2,628
		13	2,757	13	2,698
				14	2,654
Total	27,605		35,910		37,822

*During the 1979-80 assessment, several age 13 art items were included in booklet 12, which was taken by 2,748 respondents.
+There were no art exercises included in booklets 1 through 8 at age 9, booklets 1 through 9 at age 13 and booklets 1 through 8 and 10 at age 17.
++Only exercise 6 of booklet 9 at age 17 was an art exercise; it is a combined art/writing exercise.

---

each respondent. Students' names were listed with their booklet identification number so that scoring and processing personnel could go back to the school lists for data verification -- for instance, on background information -- if necessary. These lists did not leave the schools and were destroyed six months following the assessment in a school.

To provide information on respondents' backgrounds, school officials were asked to respond to a "principal's questionnaire," which included questions about the size and type of community served by the schools. In addition, in 1978-79 officials in schools were asked to respond to a "supplementary principal's questionnaire," which asked about art, music and writing/language arts programs in the school. Students also provided information on their backgrounds through questions

included in the item booklets. Samples of forms used to collect background information from students and school officials in the 1978-79 assessment appear in Appendix B.

The assessment administrator coded each student's birth date, sex, grade and racial/ethnic classification on his or her booklet. Identification numbers were precoded by machine on the front of each booklet. Administrators made a visual racial/ethnic identification at the time each student turned in his or her booklet. During the 1978-79 assessment, six different racial classifications were used: white, black, Spanish heritage, American Indian or Alaskan native, Pacific Islander or Asian, and unclassified. If an administrator was unsure of a student's racial/ethnic group, he or she referred to the student's name or questioned school personnel to make the identification. In a few cases, listening to the student's speech was an aid. The assessment administrator did not ask students to give a racial identification for themselves; however, in 1978-79 17-year-old students were asked to provide this information in one of the background questions included in the exercise booklet.

Sample sizes of the two classifications American Indian or Alaskan native and Pacific Islander or Asian are too small to permit reporting for these groups. Results for the group classified as Spanish heritage cannot be reported for separate exercises but can be reported for aggregate results across a number of exercises.

Following data collection, assessment administrators sent completed booklets to the scoring contractor, Westinghouse DataScore Systems, Iowa City, Iowa. Booklets were quality-checked to verify that correct administrative procedures were followed by the field staff. Coded identification information was also checked for accuracy; inconsistencies that could not be reconciled were sent back to the assessment administrator to be checked against the list of student names and identification numbers retained by the school for six months following the assessment.

In 1974-75, 17-year-olds who were not currently attending school were included in the assessment. These out-of-school 17-year-olds could each answer up to four booklets of assessment materials; they were paid $5.00 for each booklet they completed. Unpaced audio tapes were used for this group.

CHAPTER 6

SCORING

Scoring and computer recording of data were contracted to Westinghouse DataScore Systems, Iowa City, Iowa. While the majority of exercises in the 1978-79 art assessment were multiple-choice, several exercises were open-ended. Responses to multiple-choice exercises were read directly from the booklets by optical scanning machines. The scoring contractor employed a special staff to hand score the open-ended exercises. Scorers were responsible for categorizing responses, using the scoring guides for open-ended exercises that defined categories of acceptable and unacceptable responses. They then coded this information into ovals that could be read by optical scanning machines.

Because of the complexity and expense of art scoring, the six open-ended art exercises scheduled to be scored following the second (1978-79) art assessment were not scored until fall 1980. All of these exercises had been used in the 1974-75 assessment as well as during the 1978-79 assessment. Responses from the first assessment had been held unscored so that responses from both assessments could be scored at the same time by the same scorers. Similarly, some open-ended unreleased items from the 1978-79 assessment remained unscored so that responses can be scored with responses from a subsequent assessment.

The art exercises were scored using scoring guides that had been created by art consultants, using as a reference field trial responses from the field trials completed prior to the 1974-75 assessment. These guides were edited and revised by National Assessment staff. In September 1980, staff from NAEP and Westinghouse DataScore Systems worked to refine these scoring guides through trial use with a sample of responses to each exercise from the 1978-79 assessment. The Westinghouse staff included an artist retained as a consultant for this scoring project and several experienced scorers. Following refinement of the six scoring guides, this team of scorers assigned scores to a set of training papers. On more than one occasion, art consultants who had worked with the earlier versions of the guides were contacted for clarification.

Of the six open-ended exercises scored, two exercises required short, written an-

swers and were administered to all three ages. Four items asked for production of some kind of drawing. Two of these "drawing" items were administered to all three ages; two were administered only to 13- and 17-year-olds.

Items were scored by eight scorers organized into two teams of four. In addition, the artist consultant worked with both teams as the art authority on scoring questions and did some scoring. Each scoring team was responsible for the scoring of one short-answer and two art production exercises. Thus, each scorer had to master only three scoring guides. Since the scoring guides varied from about 15 pages for the short-answer exercises to over 100 pages for the art production exercises, it was felt that more than three guides would be excessively difficult to learn.

During training of the scorers, conducted in October 1980, scoring guides were discussed, and scorers then began working with small sets of training papers. Their scores and the master scores assigned during scoring guide refinement were compared and discrepancies were discussed. In a number of cases, inconsistencies in the master scores were detected and corrected. These discussions led to some further scoring guide revisions and refinements.

When scorers and trainers were satisfied that the scoring guides could be consistently applied, the scoring of the actual responses from the 1974-75 and 1978-79 assessments began. To help ensure that the guides were uniformly used across years and across ages, scoring of responses from all three ages and from both assessments proceeded simultaneously.

To be sure that the art hand scoring was reliable across the scoring period and across scorers, National Assessment asked Westinghouse DataScore Systems to perform two quality-control studies. In one study, a sample of about 3% of the responses to one of the six open-ended questions was drawn during each week of scoring. Each of these responses was scored by two scorers who were normally assigned to score the item. The two scorers independently rated each response and their scores were then compared. As seen in Table 6, the overall pairwise percentage of agreement was 94.6% -- Team 1 averaged 94.9% and Team 2 averaged 94.4%.

In the second quality-control study, which began near the beginning of the scoring period, scorers scored a sample set of responses; they then rescored the same set of exercises at the conclusion of the scoring period. The two scores were compared and pairwise percentages of agreement were calculated, this time within scorer and across time. Team 1 averaged 92.5% and Team 2 averaged 95.2% agreement (see Table 7).

TABLE 6. Between-Scorer Quality-Control Summaries for Art Open-Ended Scoring Summarized Across Seven Weeks

Exercise #	# of Scores	# of Agreements	% of Agreement
Team 1			
101005	2,520	2,399	95.2
301008	3,240	3,046	94.0
303042	4,800	4,574	95.2
Total	10,560	10,019	94.9
Team 2			
302006	14,280	13,566	95.0
304014	6,900	6,521	94.5
501012	1,680	1,495	89.0
Total	22,860	21,582	94.4
Grand total	33,420	31,601	94.6

TABLE 7. Within-Scorer Quality-Control Summaries for Art Open-Ended Scoring (Readings at Beginning and Conclusion of Scoring)

Exercise #	# of Scores	# of Agreements	% of Agreement
Team 1			
101005	720	654	90.8
301008	1,080	987	91.4
303042	1,680	1,577	93.9
Total	3,480	3,218	92.5
Team 2			
302006	3,060	2,939	96.0
304014	2,070	1,961	94.7
501012	360	327	90.8
Total	5,490	5,227	95.2
Grand total	8,970	8,445	94.1

The results of the ongoing quality-control study were closely monitored by Westinghouse. When necessary, corrective actions were taken, including retraining and some rescoring of work units already completed.

CHAPTER 7

DATA ANALYSIS

Measures of Achievement

The basic measure of achievement reported by National Assessment is the percentage of students responding acceptably to a given item. This percentage is an estimate of the percentage of 9-, 13- or 17-year-olds who would respond acceptably to a given item if every 9-, 13- or 17-year-old in the country were assessed.

Percentages of acceptable responses are used because each item is designed as a separate measure of some aspect of an objective or subobjective. The purpose of National Assessment is to discover if more or fewer people are able to answer these items acceptably -- and thus meet the objectives -- over time.

In addition to providing results on individual items, National Assessment reports the average performance across groups of similar items, for example, the learning area as a whole, for particular objectives or subobjectives, and so forth. The results are the mean, or arithmetic average, of the estimates of performance on the group of items, which is called the mean percentage acceptable.[1] The exercises included in the calculation of a mean percentage are usually located in several exercise booklets, and the same booklets are not administered to all students.

To present a general picture of changes in achievement,

---

[1]Twenty-two empirical distributions of change measures from the 1969-70 and 1972-73 science assessments were used to generate Monte Carlo simulations of sampling distributions for several measures of central location. In addition to the mean and median, other measures of central location that were considered in the simulation studies included the average of the extremes, two forms of biweighted estimates and three forms of weight-matching estimation described by John W. Tukey in the research paper "Some Considerations on Locators Apt for Some Squeezed-Tail (and Stretched-Tail) Parents" (1975). In almost every case, the sampling stability of the mean change was as good as or better than that of the other measures studied.

National Assessment describes the gains and losses on a group of exercises in terms of the differences in the average percentages of acceptable responses.

Unless the exercises summarized in the mean percentages of acceptable responses are identical, the means of one age group should not be compared with the means of another, since their values reflect both the choice of exercises and the performance of the students. When only a few exercises are summarized by a mean, one should be especially cautious in interpreting results, since a small set of exercises might not adequately cover the wide range of potential behaviors included under a given objective or subobjective. The mean should be interpreted literally as the arithmetic average of the percentage of acceptable responses obtained from National Assessment samples on a specific set of exercises.

In addition to providing national results, NAEP reports the achievement of various subpopulations of interest. Groups are defined by region of the country, sex, race/ethnicity, size of community, type of community, grade and level of parents' education. Results from some additional variables are also analyzed. The definitions of reporting groups are found in Appendix A. Forms used to obtain background information are presented in Appendix B.

In considering National Assessment's achievement measures, it is the differences in performance between assessments, among groups and among ages that are the most useful. By maintaining the same item or set of items in making these comparisons, we have a reasonable indicator of whether more or fewer people know or can do something judged important.

Procedures for estimating percentages of acceptable responses to exercises are dependent on the sample design. Each response by an individual is weighted and multiplied by an adjustment factor for nonresponse.[2] An estimate of the percentage of the group that would have responded to an exercise acceptably if the entire age group were assessed is defined as the weighted number of acceptable responses divided by the weighted number of all the responses. A similar ratio of weights is used to estimate percentages of acceptable responses for reporting groups or subpopulations of interest.[3]

---

[2]Appendix C discusses nonresponse in assessment samples.

[3]A weighting-class adjustment procedure was used to smooth estimated population proportions across the assessments conducted between 1970-71 and 1979-80. Appendix E discusses this procedure.

Estimating Variability in Achievement Measures

National Assessment uses a national probability sample at each age level to estimate the proportion of people who would complete an exercise in a certain way. The particular sample selected is one of a large number of all possible samples of the same size that could have been selected with the same sample design. Since an achievement measure computed from each of the possible samples would differ from one sample to another, the standard error of this statistic is used as a measure of the sampling variability among achievement measures from all possible samples. A standard error, based on one particular sample, serves to estimate that sampling variability.

In the interest of sampling and cost efficiencies, National Assessment uses a complex, stratified, multistage probability sample design. Typically, complex designs do not provide for unbiased or simple computation of sampling errors. A reasonably good approximation of standard error estimates of acceptable response percentages is obtained by applying the jackknife procedure (Miller, 1964, pp. 1594-1705; Miller, 1968, pp. 567-582; Mosteller and Tukey, 1968) to first-stage sampling units within strata. Standard errors for achievement measures such as group differences, mean percentages or mean group differences for a particular assessment year are estimated directly, taking advantage of features of the jackknife procedures that are generic to all of these statistics.[4] Since samples for different assessments are independent, the standard errors of the differences in achievement measures between assessments can be estimated simply by the square root of the sum of squared standard errors from each of the assessments.

The standard error provides an estimate of sampling reliability for the achievement measures used by National Assessment. It is comprised of sampling error and other random error associated with the assessment of a specific item or set of items. Random error includes all possible nonsystematic error associated with administering specific exercises to specific students in specific situations. For open-ended items, random differences among scorers are also included in the standard errors.

National Assessment has adhered to a standard convention whereby differences between statistics are designated as statistically significant at the .05 level of significance. That is, differences in performance between assessed years or between a reporting group and the nation are high-

---

[4]See Appendix D for a more detailed description of National Assessment's computation of standard errors.

lighted with asterisks only if they are at least twice as large as their standard error. Differences this large would occur by chance in fewer than 5% of all possible replications of the sampling and data collection procedures for any particular reporting group or national estimates.

Controlling Nonrandom Errors

Systematic errors can be introduced at any stage of an assessment -- exercise development, preparation of exercise booklets, design or administrative procedures, field administration, scoring or analysis. These nonsampling, nonrandom errors rarely can be quantified, nor can the magnitude of the bias they introduce into the estimates be evaluated directly.

Systematic errors can be controlled in large part by employing uniform administrative and scoring procedures and by requiring rigorous quality control in all phases of an assessment. If the systematic errors are the same from age to age or group to group, then the difference in percentages or mean percentages are measures with reduced bias because subtraction tends to cancel the effect of the systematic errors.

Similarly, the effect of systematic errors in different assessment years can be controlled by carefully replicating in the second assessment the procedures carried out in the first. Differences in achievement across assessment years will also be measures with reduced bias since subtraction will again tend to cancel systematic errors.

Although it is not possible for every condition or procedure to remain exactly the same between assessments conducted several years apart, National Assessment makes every effort to keep conditions as nearly the same as possible. Changes in procedures described in this report were judged to have a relatively minor impact.

APPENDIX A

DEFINITIONS OF NATIONAL ASSESSMENT REPORTING GROUPS

In addition to reporting results for all 9-, 13- and 17-year-old students in the United States, National Assessment reports results for a number of population subgroups. These subgroups are defined for both the 1974-75 and the 1978-79 assessments.

Group Definitions

Definitions of the subgroups follow:

Region

The country has been divided into four regions: Northeast, Southeast, Central and West. States included in each region are shown on the following map.

Sex

Results are reported for males and females.

Race/Ethnicity

Results are presented for blacks, whites and Hispanos. (Because of a small sample size, only average results for Hispanos can be reported by National Assessment.)

Level of Parental Education

National Assessment defines three categories of parental-education levels, based on students' reports. These categories are: (1) those whose parents did not graduate from high school, (2) those who have at least one parent who graduated from high school and (3) those who have at least one parent who has had some post-high-school education.

Type of Community

Communities in this category are defined by an occupational profile of the area served by a school as well as by the size of the community in which the school is located. This

is the only reporting group that excludes a large number of respondents. About two-thirds do not fall into the classifications listed below. Results for the remaining two-thirds are not reported since their performance is similar to that of the nation.

Advantaged-urban (high-metro) communities. Students in this group attend schools in or around cities having a population greater than 200,000 where a high proportion of the residents are in professional or managerial positions.

Disadvantaged-urban (low-metro) communities. Students in this group attend schools in or around cities having a population greater than 200,000 where a relatively high proportion of the residents are on welfare or are not regularly employed.

Rural communities. Students in this group attend schools in areas with a population under 10,000 where many of the residents are farmers or farm workers.

Size of Community

Big cities. Students in this group attend schools within the city limits of cities having a 1970 census population over 200,000.

Fringes around big cities. Students in this group attend schools within metropolitan areas (1970 U.S. Bureau of the Census urbanized areas) served by cities having a population greater than 200,000 but outside the city limits.

Medium cities. Students in this group attend schools in cities having a population between 25,000 and 200,000, not classified in the fringes-around-big-cities category.

Small places. Students in this group attend schools in communities having a population less than 25,000, not classified in the fringes-around-big-cities category.

Grade in School

Results are categorized for 9-year-olds in grades 3 and 4; 13-year-olds in grades 7 and 8; and 17-year-olds in grades 10, 11 and 12.

Modal Grade by Region

Results are categorized for 9-, 13- and 17-year-old respondents in grades 4, 8 and 11, respectively, who live in the Northeastern, Southeastern, Central or Western region of the country.

Modal Grade by Community Size

Results are categorized for 9-, 13- and 17-year-old respondents in grades 4, 8 and 11, respectively, who live in big cities, fringes around big cities, medium cities and small places.

Modal Grade by Sex

Results are categorized for 9-, 13- and 17-year-old males and females in grades 4, 8 and 11, respectively.

Visit Art Museums

Results are categorized for 9-, 13- and 17-year-old respondents who reported never having visited an art museum, having visited an art museum once and having visited an art museum five or more times.

Art Taught

Results are categorized for 9-, 13- and 17-year-olds whose principals reported that their schools offered at least one art class and for those whose principals reported that their schools offered no art classes.

Do You Collect Art

Age 13 results are categorized by those who reported collecting none, one type and two types of art. Age 17 results are categorized by those who reported collecting none, one type, two types and three types of art.

What Kinds of Artwork Do You Do

Age 9 results are categorized by those who reported doing none, 1 or 2, and 3 or 4 types of art outside of school. Age 13 and 17 results are categorized by those who reported doing none, 1 or 2, 3 or 4, and from 5 to 10 types of art outside of school.

Art Classes Taken

Age 13 results are categorized by those who reported taking none, one or two art classes. Age 17 results are categorized by those who reported taking none, one, two or three, or four to six art classes.

APPENDIX B

FORMS USED TO OBTAIN BACKGROUND INFORMATION

This appendix contains the forms used by National Assessment to collect background information from school officials and respondents in the 1978-79 art assessment. Following is a listing and a brief description of the forms included.

p. 33 School Principal's Questionnaire -- filled out by school principals or other school officials for schools at each of the age levels discussed.

p. 35 Supplementary Principal's Questionnaire (Age 9) -- given to principals of schools in which the age 9 assessment took place. The questionnaire provides information about the art and music programs of the schools.

p. 38 Supplementary Principal's Questionnaire (Age 13)[1] -- given to principals of schools in which the age 13 assessments took place. The questionnaire provides information about the art, music and writing or language arts programs of the schools.

p. 40 Supplementary Principal's Questionnaire (Age 17) -- given to principals of schools in which the age 17 assessments took place. The questionnaire provides information about the art, music and writing or language arts programs of the schools.

p. 42 Standard Background Information Form for 9-Year-Olds -- provides information about reading material in the home and level of parents' education.

---

[1]This principal's questionnaire was not used with the age 13 art exercises given in 1979-80.

p. 43 Standard Background Information Form for 13-Year-Olds -- provides information about reading material in the home, level of parents' education and place lived in at age 9.

p. 44 Standard Background Information Form for 17-Year-Olds -- provides information on homework, TV watching, racial identification, possessions in the home and classroom activities, in addition to questions also asked of 9- and 13-year-olds.

p. 47 Background Information: Art Museum or Gallery Visits -- 9-, 13- and 17-Year-Olds. This exercise asked for responses about the number of art museums or galleries visited. It was included in all 9-, 13- and 17-year-olds' art booklets given in 1978-79 except booklet 9 at age 17.[2] It was not included in the age 13 art materials assessed in 1979-80.

p. 48 Background Information: Kinds of Artwork Done Outside of School -- 9-, 13- and 17-Year-Olds. These two excercises, one for age 9 and one for ages 13 and 17, asked respondents if they had done a number of different types of artwork outside of school. The appropriate versions appear in all 1978-79 and 1979-80 booklets except booklet 9 at age 17.[2]

p. 50 Background Information: Art Classes Taken -- 13- and 17-Year-Olds. These two exercises, one for each age, asked about art classes taken or presently enrolled in. The appropriate version was included in all 1978-79 and 1979-80 booklets for 13- and 17-year olds except booklet 9 at age 17.[2]

p. 52 Background Information: Types of Art Collected -- 13- and 17-Year-Olds. There were two versions of this exercise, one for 13-year-olds and another for 17-year-olds. The appropriate version of the exercise appeared in all 1978-79 art booklets for 13- and 17-year-olds except booklet 9 at age 17.[2] This exercise was not included with the 1979-80 age 13 art materials.

---

[2]Booklet 9 for 17-year-olds included only one exercise in art. This item was an essay item designed to measure both art and writing.

School Principal's Questionnaire

This report is authorized by law (20 U.S.C. 1221 c-1). While you are not required to respond, your cooperation is needed to make the results of this survey comprehensive, accurate, and timely.

Primary Sampling Unit [ ][ ] School Number [ ][ ][ ][ ][ ][ ][ ][ ][ ]

Age Group(s) 9 13 17

PLEASE PRINT

Name of School _____

Address of School _____
(Street)

_____
(City) (State) (Zip Code)

Name of School Principal _____

Name and title of person completing the form if other than school principal

Name _____ Title _____

1. What is your best estimate of the current enrollment and the average daily attendance by grade of your school (1978-79 school year)? (Enter zeros for grades not served by your school.)

Grade	K	1	2	3	4	5	6	7	8	9	10	11	12
Enrollment													
Average Daily Attendance													

2. Approximately what percentage of the students attending your school live in each of the following areas?

_____% A In a rural area (less than 2,500)

_____% B In a town of 2,500 to 10,000

_____% C In a town of 10,000 or more

_____
100% (Items A-C should add to 100%)

BEST COPY AVAILABLE

3. Approximately what percentage of the students attending your school are children of

_____% A Professional or managerial personnel

_____% B Sales, clerical, technical or skilled workers

_____% C Factory or other blue collar workers

_____% D Farm workers

_____% E Persons not regularly employed

_____% F Persons on welfare

_____ (Items A-F should add to 100%)
100%

4. Approximately what percentage of the students attending your school are

_____% A American Indian or Alaskan Native

_____% B Asian or Pacific Islander

_____% C Hispanic, regardless of race (Mexican, Puerto Rican, Cuban, Central or South American or other Spanish culture or orig.)

_____% D Black and not Hispanic

_____% E White and not Hispanic

_____ (Items A-E should add to 100%)
100%

5. Does your school qualify for ESEA Title I assistance?

_____ Yes - If Yes, approximately what number of students qualify for and what number of students are receiving ESEA Title I assistance?

_____ Approximate number of students qualifying for ESEA Title I assistance

_____ Approximate number of students receiving ESEA Title I assistance

_____ No

THANK YOU FOR YOUR COOPERATION

Supplementary Principal's Questionnaire (Age 9)

Instructions: The purpose of this questionnaire is to provide additional information which will be used in the analyses of NAEP data. Darken the appropriate ovals with a soft lead pencil. If you have questions about any of the following items, please contact the National Assessment district supervisor. Thank you for your cooperation.

1. Is art taught in your school?

   ⬭ Yes ⬭ No (Go to Question 10 on page 3)

2. Are there specified and systematically ordered objectives for art instruction in your school?

   ⬭ Yes ⬭ No

3. Which of the following items are available as art teaching aids in your school?

A. Slides of works of art	Yes ⬭	No ⬭
B. Films on art	Yes ⬭	No ⬭
C. Film strips of works of art	Yes ⬭	No ⬭
D. Color reproductions of works of art	Yes ⬭	No ⬭
E. Art books	Yes ⬭	No ⬭
F. Original works of art	Yes ⬭	No ⬭
G. Other	Yes ⬭	No ⬭

4. Does your school have a special art room?

   ⬭ Yes ⬭ No

5. Which of the following art equipment is found in your school?

A.	Pottery kiln	Yes ⬭	No ⬭
B.	Potters wheel	Yes ⬭	No ⬭
C.	Weaving loom	Yes ⬭	No ⬭
D.	Easels	Yes ⬭	No ⬭
E.	Cameras	Yes ⬭	No ⬭
F	Other	Yes ⬭	No ⬭

6. Do the nine-year-old students in your school receive art instruction?

⬭ Yes ⬭ No (Go to Question 10 on page 3)

7. If the nine-year-old students receive art instruction, who does the teaching? (Mark each appropriate oval.)

A.	The classroom teacher	Yes ⬭	No ⬭
B.	A member of a teaching team having a background in art but not certified in art	Yes ⬭	No ⬭
C.	A certified art teacher	Yes ⬭	No ⬭
D.	Other	Yes ⬭	No ⬭

8. On the average, how much time each week is given to art instruction of nine-year-old students?

⬭ Less than 30 minutes
⬭ 30 to 59 minutes
⬭ 60 to 89 minutes
⬭ 90 to 119 minutes
⬭ 120 minutes and over

9. For nine-year-old students, are field trips taken to art museums or art galleries?

⬭ Yes ⬭ No

Supplementary Principal's Questionnaire (Age 13)

Instructions: The purpose of this questionnaire is to provide additional information which will be used in the analyses of NAEP data. Darken the appropriate ovals with a soft lead pencil. If you have questions about any of the following items, please contact the National Assessment District Supervisor. Thank you for your cooperation.

1. Is art taught in your school?

   ⬭ Yes ⬭ No (Go to Question 12 on page 3)

2. Are there specified and systematically ordered objectives for art instruction in your school?

   ⬭ Yes ⬭ No

3. Which of the following items are available as art teaching aids in your school?

A.	Slides of works of art	Yes ⬭	No ⬭
B.	Films on art	Yes ⬭	No ⬭
C.	Film strips of works of art	Yes ⬭	No ⬭
D.	Color reproductions of works of art	Yes ⬭	No ⬭
E.	Art books	Yes ⬭	No ⬭
F.	Original works of art	Yes ⬭	No ⬭
G.	Other	Yes ⬭	No ⬭

4. Are students <u>required</u> to take art in seventh grade?

   ⬭ Yes ⬭ No (Go to Question 6)

   ⬭ No seventh grade (Go to Question 8)

5. If there is a requirement, for what length of time do the students take art?

   ⬭ 1/4 school year or less

   ⬭ 1/2 school year

   ⬭ 3/4 school year

   ⬭ Full school year

6. Is an <u>elective</u> art course available for seventh grade students?

   ⬭ Yes ⬭ No (Go to Question 8)

7. If you said "Yes" to Question 6, approximately what percentage of seventh graders elect to take art courses each year?

- ⬭ 0 to 24%
- ⬭ 25 to 49%
- ⬭ 50 to 74%
- ⬭ 75 to 89%
- ⬭ 90 to 100%

8. Are the students required to take art in eighth grade?

- ⬭ Yes ⬭ No (Go to Question 10 on page 3)
- ⬭ No eighth grade (Go to Question 12 on page 3)

9. If there is a requirement, for what length of time do the students take art?

- ⬭ 1/4 school year or less
- ⬭ 1/2 school year
- ⬭ 3/4 school year
- ⬭ Full school year

10. Is an elective art course available for eighth grade students?

- ⬭ Yes ⬭ No (Go to Question 12)

11. If you said "Yes" to Question 10, approximately what percentage of eighth graders elect to take art courses each year?

- ⬭ 0 to 24%
- ⬭ 25 to 49%
- ⬭ 50 to 74%
- ⬭ 75 to 89%
- ⬭ 90 to 100%

Instructions: The purpose of this questionnaire is to provide additional information which will be used in the analyses of NAEP data. Darken the appropriate ovals with a soft lead pencil. If you have questions about any of the following items, please contact the National Assessment district supervisor. Thank you for your cooperation.

1. Is art taught in your school?

   ⬭ Yes ⬭ No (Go to Question 7 on page 3)

2. Are there specified and systematically ordered objectives for art instruction in your school?

   ⬭ Yes ⬭ No

3. Which of the following items are available as art teaching aids in your school?

A.	Slides of works of art	Yes ⬭	No ⬭
B.	Films on art	Yes ⬭	No ⬭
C.	Film strips of works of art	Yes ⬭	No ⬭
D.	Color reproductions of works of art	Yes ⬭	No ⬭
E.	Art books	Yes ⬭	No ⬭
F.	Original works of art	Yes ⬭	No ⬭
G.	Other	Yes ⬭	No ⬭

4. Is art instruction available in your school for the following grades?

		No such Grade	Yes	No
A.	Grade 9	⬭	⬭	⬭
B.	Grade 10	⬭	⬭	⬭
C.	Grade 11	⬭	⬭	⬭
D.	Grade 12	⬭	⬭	⬭
E.	Ungraded	⬭	⬭	⬭

5. Does your school offer the following types of art instruction for Grades 9, 10, 11 or 12?

	Yes	No
A. General art	⬭	⬭
B. Fine art	⬭	⬭
C. Crafts	⬭	⬭
D. Humanities	⬭	⬭
E. Art history	⬭	⬭
F. Commercial art	⬭	⬭
G. Other	⬭	⬭

6. About what percent of the 17-year-old students are currently enrolled in art classes?

⬭ 0 to 24%

⬭ 25 to 49%

⬭ 50 to 74%

⬭ 75 to 89%

⬭ 90 to 100%

BEST COPY AVAILABLE

1. Does your family get a newspaper regularly?

   ⬭ Yes ⬭ No ⬭ I don't know.

2. Does your family get any magazines regularly?

   ⬭ Yes ⬭ No ⬭ I don't know.

3. Are there more than 25 books in your home?

   ⬭ Yes ⬭ No ⬭ I don't know.

4. Is there an encyclopedia in your home?

   ⬭ Yes ⬭ No ⬭ I don't know.

5. How much school did your father complete?
   (FILL IN THE ONE OVAL which best shows how much school your father completed.)

   ⬭ Did not complete the 8th grade

   ⬭ Completed the 8th grade, but did not go to high school

   ⬭ Went to high school, but did not graduate from high school

   ⬭ Graduated from high school

   ⬭ Some education after graduation from high school

   ⬭ I don't know.

6. Did your father graduate from a college or university?

   ⬭ Yes ⬭ No ⬭ I don't know.

7. How much school did your mother complete?
   (FILL IN THE ONE OVAL which best shows how much school your mother completed.)

   ⬭ Did not complete the 8th grade

   ⬭ Completed the 8th grade, but did not go to high school

   ⬭ Went to high school, but did not graduate from high school

   ⬭ Graduated from high school

   ⬭ Some education after graduation from high school

   ⬭ I don't know.

8. Did your mother graduate from a college or university?

   ⬭ Yes ⬭ No ⬭ I don't know.

1. Does your family get a newspaper regularly?

   ⬭ Yes ⬭ No ⬭ I don't know.

2. Does your family get any magazines regularly?

   ⬭ Yes ⬭ No ⬭ I don't know.

3. Are there more than 25 books in your home?

   ⬭ Yes ⬭ No ⬭ I don't know.

4. Is there an encyclopedia in your home?

   ⬭ Yes ⬭ No ⬭ I don't know.

5. How much school did your father complete?
   (FILL IN THE ONE OVAL which best shows how much school your father completed.)
   ⬭ Did not complete the 8th grade
   ⬭ Completed the 8th grade, but did not go to high school
   ⬭ Went to high school, but did not graduate from high school
   ⬭ Graduated from high school
   ⬭ Some education after graduation from high school

   ⬭ I don't know.

6. Did your father graduate from a college or university?

   ⬭ Yes ⬭ No ⬭ I don't know.

7. How much school did your mother complete?
   (FILL IN THE ONE OVAL which best shows how much school your mother completed.)
   ⬭ Did not complete the 8th grade
   ⬭ Completed the 8th grade, but did not go to high school
   ⬭ Went to high school, but did not graduate from high school
   ⬭ Graduated from high school
   ⬭ Some education after graduation from high school

   ⬭ I don't know.

8. Did your mother graduate from a college or university?

   ⬭ Yes ⬭ No ⬭ I don't know.

9. Where did you live on your ninth birthday?
   ⬭ In the United States (Please specify the state or territory.)

   _____

   ⬭ Outside the United States (Please specify the country.)

   _____

   ⬭ I don't know.

Standard Background Information Form for 17-Year-Olds

1. Which of the following does your family have at home? (Fill in one oval on each line.)

		Have	Do not have
A.	Newspaper received regularly	⬭	⬭
B.	Magazines received regularly	⬭	⬭
C.	More than 25 books	⬭	⬭
D.	Encyclopedia	⬭	⬭
E.	Dictionary	⬭	⬭
F.	Record player	⬭	⬭
G.	Tape recorder or cassette player	⬭	⬭
H.	Typewriter	⬭	⬭
I.	Vacuum cleaner	⬭	⬭
J.	Electric dishwasher	⬭	⬭
K.	Two or more cars or trucks that run	⬭	⬭

2. How much time did you spend on homework yesterday?

⬭ No homework was assigned
⬭ I had homework but didn't do it
⬭ Less than one hour
⬭ Between 1 and 2 hours
⬭ More than 2 hours

3. How many different schools have you attended since you started the first grade?

⬭ 1 to 3 schools
⬭ 4 to 6 schools
⬭ 7 to 9 schools
⬭ 10 or more schools

4. How long have you lived in the community in which you now live?

⬭ All my life
⬭ 10 or more years but not all my life
⬭ 5 to 9 years
⬭ 2 to 4 years
⬭ 1 year
⬭ Less than 1 year

5. How much television did you watch yesterday?

⬭ None
⬭ 1 hour or less
⬭ 1 hour
⬭ 2 hours
⬭ 3 hours
⬭ 4 hours
⬭ 5 hours
⬭ 6 hours or more

6. Is English the language spoken most often in your home?

⬭ Yes ⬭ No

7. Is a language other than English spoken in your home?

⬭ Often ⬭ Sometimes ⬭ Never

8. How many brothers or sisters do you have who are <u>older</u> than you?

None ⬭ 1 ⬭ 2 ⬭ 3 ⬭ 4 ⬭ 5 ⬭ 6 or more ⬭

9. How many brothers or sisters do you have who are <u>younger</u> than you?

None ⬭ 1 ⬭ 2 ⬭ 3 ⬭ 4 ⬭ 5 ⬭ 6 or more ⬭

10. A. What is your racial background?

⬭ American Indian or Alaskan Native
⬭ Asian or Pacific Islander
⬭ Black
⬭ White
⬭ Other (Please specify) _____

B. Is your ethnic heritage Hispanic (such as Mexican, Puerto Rican, Cuban, Central or South American or other Spanish Culture or origin)?

⬭ Yes ⬭ No

11. How often has each of the following been used in the courses you are taking this year? (Fill in one oval on each line.)

		Never	Seldom	Fairly Often	Frequently
A.	Lirtening to the teacher's lecture	⬭	⬭	⬭	⬭
B.	Participating in student-centered discussions	⬭	⬭	⬭	⬭
C.	Working on a project or in a laboratory	⬭	⬭	⬭	⬭
D.	Writing essays, themes, poetry, stories	⬭	⬭	⬭	⬭
E.	Going on field trips	⬭	⬭	⬭	⬭
F.	Having individualized instruction (small groups or one-to-one with a teacher)	⬭	⬭	⬭	⬭
G.	Using teaching machines or computer-assisted instruction	⬭	⬭	⬭	⬭
H.	Watching television lectures	⬭	⬭	⬭	⬭
I.	Studying from textbooks	⬭	⬭	⬭	⬭
J.	Library or media center assignments	⬭	⬭	⬭	⬭

A. How much school did your father complete?
(FILL IN THE <u>ONE</u> OVAL which best shows how much school your father completed.)

⬭ Did not complete the 8th grade
⬭ Completed the 8th grade, but did not go to high school
⬭ Went to high school, but did not graduate from high school
⬭ Graduated from high school
⬭ Some education after graduation from high school

⬭ I don't know.

B. Did your father graduate from a college or university?

⬭ Yes ⬭ No ⬭ I don't know.

C. How much school did your mother complete?
(FILL IN THE ONE OVAL which best shows how much school your mother completed.)

⬭ Did not complete the 8th grade
⬭ Completed the 8th grade, but did not go to high school
⬭ Went to high school, but did not graduate from high school
⬭ Graduated from high school
⬭ Some education after graduation from high school

⬭ I don't know.

D. Did your mother graduate from a college or university?

⬭ Yes ⬭ No ⬭ I don't know.

E. Where did you live on your ninth birthday?

⬭ In the United States (Please specify the state or territory.)

_____

⬭ Outside of the United States (Please specify the country.)

_____

⬭ I don't know.

F. Where did you live on your thirteenth birthday?

⬭ In the United States (Please specify the state or territory.)

_____

⬭ Outside of the United States (Please specify the country.)

_____

⬭ I don't know.

Background Information: Art Museum or Gallery Visits
9-, 13- and 17-Year-Olds

The works of artists are shown in art museums and art galleries. How often have you visited art museums or art galleries?

- ⬭ **Never**
- ⬭ **One time**
- ⬭ **About five times**
- ⬭ **About ten times**
- ⬭ **Fifteen or more times**

Background Information: Kinds of Artwork Done Outside of School
9-Year-Olds

Outside of school, what kinds of art work do you do? Tell whether you do each of the following things.

A. Outside of school, do you draw?

- ⊂ Yes
- ⊂ No
- ⊂ I don't know.

B. Outside of school, do you paint?

- ⊂ Yes
- ⊂ No
- ⊂ I don't know.

C. Outside of school, do you make collages by cutting and pasting paper, cloth and scrap materials?

- ⊂ Yes
- ⊂ No
- ⊂ I don't know.

D. Outside of school, do you carve or make models with wood, stone, clay, metal or plastic?

- ⊂ Yes
- ⊂ No
- ⊂ I don't know.

Background Information: Kinds of Artwork Done Outside of School
13- and 17-Year-Olds

Outside of school, what kinds of art work do you do? Tell whether you do each of the following things outside of school.

		Yes	No	I don't know.
A.	Drawing	⬭	⬭	⬭
B.	Painting	⬭	⬭	⬭
C.	Making pictures by cutting and pasting paper, cloth and scrap materials	⬭	⬭	⬭
D.	Carving or modeling in wood, stone, clay, metal or plastic	⬭	⬭	⬭
E.	Print-making such as block printing, silk screening, etching	⬭	⬭	⬭
F.	Making pottery, ceramics or mosaics	⬭	⬭	⬭
G.	Weaving, macrame or knot-tying, or needlework such as embroidery, needlepoint, knitting, crocheting	⬭	⬭	⬭
H.	Making photographs or films	⬭	⬭	⬭
I.	Making jewelry	⬭	⬭	⬭
J.	Creating designs or plans for things like clothes, toys, cars, houses, furniture	⬭	⬭	⬭

Background Information: Art Classes Taken -- 13-Year-Olds

A. Did you take an art class in school last year?

⬭ Yes ⬭ No ⬭ I don't know.

B. Are you taking an art class during this school year?

⬭ Yes ⬭ No ⬭ I don't know.

Background Information: Art Classes Taken — 17-Year-Olds

A. Please indicate in which of the following years you took art classes in school.

	Yes	No	I don't know.
Seventh grade	⬭	⬭	⬭
Eighth grade	⬭	⬭	⬭
Ninth grade	⬭	⬭	⬭
Tenth grade	⬭	⬭	⬭
Eleventh grade	⬭	⬭	⬭
Twelfth grade	⬭	⬭	⬭

B. Are you currently enrolled in an art class?

⬭ Yes ⬭ No ⬭ I don't know.

Background Information: Types of Art Collected — 13-Year-Olds

A. Original works of art are such things as paintings, drawings, sculpture, ceramic pieces, jewelry, hand-signed prints, weaving or any other work actually made by an artist.

Do you collect original works of art?

- ⊂ Yes
- ⊂ No

- ⊂ I don't know.

B. Reproductions of works of art are copies of original art works.

Do you collect reproductions of works of art?

- ⊂ Yes
- ⊂ No

- ⊂ I don't know.

Background Information: Types of Art Collected — 17-Year-Olds

A. Original works of art are such things as paintings, drawings, sculpture, ceramic pieces, jewelry, hand-signed prints, weaving or any other work actually made by an artist.

Do you collect original works of art?

- ⬭ Yes
- ⬭ No

- ⬭ I don't know.

B. Reproductions of art works are copies of original art works.

Do you collect reproductions of works of art?

- ⬭ Yes
- ⬭ No

- ⬭ I don't know.

C. Antiques are old objects such as furniture, glassware, rugs and toys.

Do you collect antiques?

- ⬭ Yes
- ⬭ No

- ⬭ I don't know.

APPENDIX C

RESPONSE RATES FOR ASSESSMENT SAMPLES

Table C-1 shows the response rates for students assessed in 1974-75 and 1978-79. In the 1974-75 assessment, for each of the three age groups, 12 students and 4 alternates were selected for each assessment session. If all 12 students appeared for the session, then the alternates were dismissed. Otherwise, enough alternates were selected to bring the size of the group up to or as near as possible to 12. If the

---

TABLE C-1. Number of Students Assessed and Percent of Sample Covered, by Age and Assessment Year

Year	Age	Number of Packages+	Total Number of Students Assessed	Average Number Assessed Per Package	Average Sample Covered in Percent
1974-75	9	6 (12)	28,403	2,367	87.5
	13	7 (13)	30,312	2,332	83.7
	17++	7 (13)	29,612	2,279	69.7
1978-79	9	11	27,605	2,510	87.8
	13	13	35,910	2,762	84.9
	17	14	37,822	2,702	77.8

+In the 1974-75 assessment, there were three reading booklets at each age that were triple sampled. This increases the effective number of booklets by six at each age. The effective number of booklets is shown in parentheses and is used to determine the average number assessed per booklet.

++At age 17, these numbers include both 17-year-olds enrolled and not enrolled in school for the reading and art assessments. In 1974-75, there were two additional booklets of experimental materials and two booklets of materials administered under a special contract. The respondents for these booklets are not included in these figures.

group assessed numbered between 8 and 12 students, then the administration was considered complete. If the final total was not at least a quorum of 8, a second and sometimes a third make-up session was held. The percentages in Table C-1 are based on the numbers of students assessed from the original groups of 12 selected and do not reflect the use of alternates.

For the 1978-79 assessment, slightly different procedures were used. The number of students selected for each administrative session varied from 16-25 students, depending on previous response rates obtained from schools in similar communities. No alternates were selected. The quorum size needed to consider an administrative session complete varied according to the number of students selected. Since nonresponse rates have always been relatively small for ages 9 and 13, the make-up or follow-up procedures used in 1978-79 for those ages were similar to the ones used for the first art assessment. If a quorum was not obtained at the first administrative session, a second and sometimes a third make-up session was held. At age 17, in the 1978-79 assessment, follow-up procedures were conducted on a school, rather than a session, basis. If a school had an overall response rate of less than 75%, then all nonrespondents in the school were contacted for one or two follow-up sessions. These follow-up procedures for 17-year-olds provided sample coverage similar to that obtained at ages 9 and 13.

Since response rates at age 17 have always been somewhat lower than at the other two ages, the Research Triangle Institute (RTI), Raleigh, North Carolina, was asked to conduct a special study of nonrespondents during the 1972-73 assessment of science and mathematics. The results (Kalsbeek et al., 1975; Rogers et al., 1977) indicated that about 80% of the total nonrespondent group did not appear at the assessment sessions because of conflicting school activities or illness. The remaining 20% did not seem to be available. They attended school infrequently, if at all (for practical purposes, they had dropped out), or they had moved out of the school attendance area. In either case, these students probably should not have been listed in the in-school population of eligibles.

Tables published in previous National Assessment reports showing response rates for age 17 generally contain percentages adjusted to account for those 17-year-olds listed, but not attending school. But, since National Assessment has not had the resources in recent assessment years to replicate the RTI study, the 20% figure used as a basis for adjusting these percentages may be outdated, and thus the percentages given in Table C-1 have not been adjusted. It seems likely that despite efforts to update the lists of

eligibles, these lists still contain some percentage of students who have in effect left the schools. Thus, the percentages listed for age 17 are probably underestimates of the actual response rates for 17-year-olds attending school.

APPENDIX D

COMPUTATION OF MEASURES OF ACHIEVEMENT, CHANGES IN ACHIEVEMENT AND STANDARD ERRORS

Several measures of achievement that National Assessment uses in its reports are described in Chapter 7 of this document. The sample design, as described in Chapter 4, is a complex, deeply stratified, multistage probability sample design. Measures of achievement are obtained by weighting individual responses appropriately. Reasonably good approximation of standard error estimates of these achievement measures can be obtained by applying the jackknife procedure to first-stage sampling units within strata, using the method of successive differences and accumulating across strata.

In this section, the measures of achievement are first defined in algebraic form, followed by a description of the jackknife method that National Assessment uses to estimate the standard errors of achievement measures.

Measures of Achievement

Based on the sample design, a weight is assigned to every individual who responds to an exercise administered in an assessment. The weight is the reciprocal of the probability of selecting a particular individual to take a particular exercise with adjustment for nonresponse. Since the probabilities of selection are based on an estimated number of people in the target age population, the weight for an individual estimates the number of similar people that individual represents in the age population. As explained in Appendix E, the weights are adjusted to reflect information from previous assessments on population distributions.

A sum of the weights for all individuals at an age level responding to an exercise is an estimate of the total number of people in that age population. A sum of weights for all individuals at an age responding correctly to an exercise is an estimate of the number of people in the age population who would be able to respond correctly if the entire population were assessed. These concepts also apply to any reporting group (e.g., defined by region, sex, and so on) and category of response (e.g., correct, incorrect and "I don't know").

Let W^{e}_{ihk} = sum of weights for respondents to exercise e who are in reporting subgroup i and who are in the kth replicate of the hth sampling stratum, and

C^{ej}_{ihk} = sum of weights for respondents to exercise e who are in reporting subgroup i, who are in the kth replicate of the hth sampling

stratum and who selected response category j (e.g., correct foil) for the exercise.

Note that $W^{e}_{ihk} = \sum_{j} C^{ej}_{ihk}$.

Then summing k over the n_h sample replicates in the stratum h, and summing over the H sampling strata,

$$W^{e}_{i++} = \sum_{h=1}^{H} \sum_{k=1}^{n_h} W^{e}_{ihk}$$

estimates the number of eligibles in the population who are in subgroup i.

Similarly, $C^{ej}_{i++} = \sum_{h=1}^{H} \sum_{k=1}^{n_h} C^{ej}_{ihk}$ estimates the number of eligibles in the population who are in subgroup i and who would select response category j for exercise e.

An estimate of the proportion of the eligibles in the age population in group i who would select response category j on exercise e is:

(1) $P^{ej}_{i} = C^{ej}_{i++}/W^{e}_{i++}$.

In the special case where the proportion of all age-eligibles who would select response category j on exercise e is estimated, the index A (for ALL) will be used in place of i as follows:

(2) $P^{ej}_{A} = C^{ej}_{A++}/W^{e}_{A++}$.

In National Assessment reports, the proportion in (1) multiplied by 100 is called the group percentage, and the proportion in (2) multiplied by 100 is called the national percentage. The difference between the proportion in subgroup i who would select category j on exercise e and the proportion in the nation is denoted by:

(3) $\Delta P^{ej}_{i} = P^{ej}_{i} - P^{ej}_{A}$.

National Assessment also reports the arithmetic mean of the percentage of correct responses over sets of exercises corresponding to the measures in (1), (2) and (3). These means are taken over the set of all exercises or a subset of exercises classified by a reporting topic or content objective. The mean percentages of correct responses taken over m exercises in some set of exercises corresponding to measures (1), (2) and (3) are, respectively:

(4) $\bar{P}_{i} = \frac{1}{m} \sum_{e} C^{e}_{i++}/W^{e}_{i++}$,

(5) $\bar{P}_{A} = \frac{1}{m} \sum_{e} C^{e}_{A++}/W^{e}_{i++}$ and

(6) $\overline{\Delta P}_{i} + \bar{P}_{i} - \bar{P}_{A}$.

Note that the response category subscript j has been suppressed since the means are understood to be taken over the correct response category for each exercise.

Each of these six achievement measures is computed and routinely used in reports describing achievement data for any assessment. The simple difference in these measures between two assessments of the same exercise (or sets of exercises) provides six measures of change in achievement that are routinely used in National Assessment's change reports. The next section describes how standard errors are estimated for the 12 statistics

used in NAEP reports.

Computation of Standard Errors

In order to obtain an approximate measure of the sampling variability in the statistics (1) through (6), a jackknife replication procedure for estimating the sampling variance of nonlinear statistics from complex, multistage samples was tailored to National Assessment's sample design. Miller (1968, 1974) and Mosteller and Tukey (1977) provide information about the jackknife technique, while Folsom (1977) describes how the procedure is used in estimating standard errors for National Assessment's sample design.

To demonstrate the computational aspects of this technique, consider estimating the variance of the statistic in (1) — the proportion of age-eligibles in subgroup i who would select response category j on exercise e.

This statistic is based on the data from all the n_h replicates in the H strata. Let P^{ej}_{i-hk} be defined as a replication estimate of P^{ej}_i and constructed from all the replicates excluding the data from replicate k in stratum h. These replication estimates are computed as if the excluded replicate had not responded, and a reasonable nonresponse adjustment is used to replace the data in replicate hk in estimating P^{ej}_i. Several choices for replacing the data in replicate hk are available. In order to obtain a convenient and computationally efficient algorithm for approximating standard errors, National Assessment replaces C^{ej}_{ihk} and W^{e}_{ihk} from the hkth replicate with corresponding sums from another paired replicate in the same stratum. The replicate estimate is then computed. The replicate estimates to be used in the calculations are determined by arranging all the replicates in each stratum into successive pairs. That is, replicate 1 is paired with replicate 2, replicate 2 with replicate 3, 3 with 4,...(n_h-1) with n_h and replicate n_h with replicate 1.

The contribution to the variance of P^{ej}_i by each pair of replicates is the change in the value of the statistic incurred by replacing the data from each replicate in the pair with the data from the other replicate in the pair and recomputing P^{ej}_i in the usual way. This produces two replicate estimates. Squaring the difference between these replicate estimates and then dividing by eight measures the contribution of this pair of replicates to the total variance. The sum of these contributions over all n_h successive pairs in the stratum is the contribution by stratum h to the total variance. The square root of the sum of the H stratum contributions is the estimate of the standard error of P^{ej}_i.

Algebraically, the two replicate estimates for the pair k, $k+1$ (where $k=1,\ldots n_h$ and $n_h+1=1$) are:

$$(7)\quad P^{ej}_{i-hk} = \frac{C^{ej}_{i++} - C^{ej}_{ihk} + C^{ej}_{ih(k+1)}}{W^{ej}_{i++} - W^{ej}_{ihk} + W^{ej}_{ih(k+1)}}$$

and

$$(8)\quad P^{ej}_{i-h(k+1)} = \frac{C^{ej}_{i++} - C^{ej}_{ih(k+1)} + C^{ej}_{ihk}}{W^{ej}_{i++} - W^{ej}_{ih(k+1)} + W^{ej}_{ihk}}.$$

The contribution to the total variance from stratum h is:

$$(9)\quad \text{var}\,(P^{ej}_{ih}) = \tfrac{1}{8} \sum_{k}^{n_h} \left(P^{ej}_{i-hk} - P^{ej}_{i-h(k+1)}\right)^2.$$

And finally, an estimate of the standard error of P^{ej}_i is:

$$(10)\quad \text{SE}\,(P^{ej}_i) = \left(\sum_{h}^{H} \text{var}\; P^{ej}_{ih}\right)^{1/2}.$$

Multiplying P^{ej}_i by 100 yields the percentage of response to category j. Multiplying $SE(P^{ej}_i)$ by 100 yields the corresponding estimated standard error of the percentage.

In general, the jackknifed standard errors of the proportion estimates will be larger than the simple random sampling formula $(PQ/N)^{1/2}$, where $P=P^{ej}_i$, $Q=1-P$ and N is the number of sampled respondents in subgroup i who took the exercise. The larger size of $SE(P^{ej}_i)$ reflects mainly the loss of precision due to cluster-sampling of schools and students. The standard errors for the achievement measures (2) through (6) are computed through a series of steps analogous to those followed in computing $SE(P^{ej}_i)$.

The standard errors for the differences between two assessments for any of the achievement measures (1) through (6) are computed as the square root of the sum of the squared standard errors from each of the separate assessments.

The size of the standard errors depends largely not only on the number of replicates and schools included in the sample, but also on the number of respondents in each of the reporting groups. Table D-1 shows the average number of students responding to an exercise booklet for each of the reporting groups for each age for each assessment year. Table D-2 shows National Assessment's current estimates of the proportions of students in each reporting group at each age.

TABLE D-1. Average Number of Respondents in Reporting Groups Taking an Item Booklet, by Age and Assessment Year+

	Age 9		Age 13		Age 17#	
	1974-75	1978-79	1974-75	1978-79	1974-75	1978-79
Nation	2,382	2,510	2,388	2,762	2,168	2,702
Region						
Northeast	581	580	601	675	536	642
Southeast	564	625	559	657	501	683
Central	695	665	686	752	634	725
West	542	639	542	681	496	649
Sex						
Male	1,197	1,255	1,172	1,370	1,050	1,312
Female	1,184	1,255	1,216	1,395	1,118	1,386
Race/ethnicity						
White	1,844	1,849	1,884	2,053	1,794	2,134
Black	396	484	374	507	279	392
Other++	142	177	130	202	96	176
Parental education						
Not graduated high school	215	191	355	326	357	386
Graduated high school	514	589	746	895	696	895
Post high school	758	880	973	1,200	993	1,334
Unknown++	895	850	314	341	122	87
Type of community						
Rural	249	252	262	271	218	246
Disadvantaged urban	236	250	227	281	228	306
Advantaged urban	238	253	236	277	202	265
Other++	1,659	1,755	1,663	1,933	1,521	1,885
Size of community						
Big cities	460	723	460	775	390	718
Fringes around big cities	450	448	472	588	423	571
Medium cities	346	237	326	287	326	287
Small places	1,126	1,097	1,131	1,115	1,030	1,123
Grade						
3, 7, 10	522	623	615	683	282	360
4, 8, 11	1,793	1,834	1,713	2,031	1,593	2,014
12	—	—	—	—	253	285
Other++	67	53	60	48	36	43

+Data may not total due to rounding error.
#Seventeen-year-olds enrolled in school.
++Data are not reported for these groups.

TABLE D-2. Estimated Current Population Proportions of National Assessment Reporting Groups for In-School Students

Reporting Groups	Age 9	Age 13	Age 17
Sex			
Male	.50	.50	.48
Female	.50	.50	.52
Race/ethnicity			
White	.79	.80	.83
Black	.14	.13	.12
Other	.07	.07	.05
Region			
Northeast	.25	.25	.25
Southeast	.22	.23	.20
Central	.27	.27	.29
West	.26	.25	.26
Parental education			
Not graduated high school	.09	.13	.15
Graduated high school	.24	.32	.32
Post high school	.33	.42	.48
Unknown	.34	.13	.05
Type of community			
Rural	.08	.10	.08
Disadvantaged urban	.07	.07	.09
Advantaged urban	.11	.11	.11
Other	.74	.72	.72
Size of community			
Big cities	.20	.21	.19
Fringes around big cities	.22	.22	.26
Medium cities	.12	.11	.11
Small places	.46	.46	.44
Grade in school			
<3, <7, <10	<.01	.02	.02
3, 7, 10	.23	.25	.13
4, 8, 11	.75	.72	.75
>4, >8, 12	<.01	<.01	.10
Other	<.01	<.01	<.01

APPENDIX E

ADJUSTMENT OF RESPONDENT WEIGHTS BY SMOOTHING TO REDUCE RANDOM VARIABILITY OF ESTIMATED POPULATION PROPORTIONS

Background

A weight is assigned to every individual who responds to an exercise administered in an assessment. The weight is the reciprocal of the probability of selection of the individual with adjustment for nonresponse. The weight for an individual estimates the number of people that the individual represents in the age population. The sum of the weights of all individuals at an age level who responded to an exercise is an estimate of the total number of people in that age population in the year that the exercise was assessed. Similarly, the sum of weights for all individuals who took the exercise and who also are members of some demographic category (such as blacks) gives an estimate of the number of people in the age population, for the year, who are also members of the category. The ratio of the two totals estimates the proportional representation of the demographic category in the age population for the given year.

Separate estimates of the proportional representation of the various demographic subgroups are provided by each booklet administered to a particular age group in a given year. Because of sampling variability, the estimates of population proportions for a given year based on single booklets vary. There is also sampling variation in estimates of population proportions from year to year in addition to any existing trends in population proportions over time.

It is desirable to reduce the random variability of population proportions as much as possible, since this variability has an effect on performance estimates. For example, the percentage of acceptable responses for an age group is a function of the relative proportions of high-performing and low-performing groups. If the relative proportions of these groups are very different in different assessments due to sampling variability, then a portion of the change in percentage of acceptable responses for an age group could be attributable to yearly sampling differences in the relative proportions of high- and low-achieving groups.

In addition to reporting performance estimates for an age

group as a whole, National Assessment also reports performance for various subpopulations, such as whites or blacks. Because variability of subgroups within these subpopulations (such as males and females within the white population) influences the performance estimates for the subpopulations, it is desirable that fluctuations of proportions of all subgroups of each subpopulation be reduced as much as possible.

For each age and year, each of the various booklets administered provides estimates of a given population proportion. Since these estimates are subject to booklet-to-booklet variability, a better estimate of the population proportion, which will have reduced variability, is obtained by combining the information from all booklets. However, these proportions vary from year to year due to random sampling variability or systematic differences in sampling procedures. An even better estimate of population proportions for any single year can be obtained by smoothing the proportions over several assessment years. The word "smoothing" is used here in the sense of fitting a smooth curve to a sequence of numbers by robust/resistant procedures (Tukey, 1977). Smoothing estimates of population proportions reduces a large portion of the sampling variability while preserving, as far as possible, actual trends occurring in the age population.

After the population proportions have been smoothed, adjusted weights are derived for the assessed individuals so that the population proportions computed using the adjusted weights are equal to the smoothed proportions. The adjusted weights are then used for all analyses.

Smoothing Procedures Used by National Assessment

The most direct way to smooth proportions is first to classify people into mutually exclusive multiway cells on the basis of their membership in categories of various important variables and then to smooth the proportions within each of the resulting multiway cells across years. Unfortunately, this procedure tends to produce a large number of cells with few people and, consequently, quite unstable estimates of smoothed proportions.

To circumvent this difficulty, National Assessment has utilized various smoothing procedures. These procedures, which are all basically weighting-class adjustments applied independently to each age, are designed to control, to varying degrees, fluctuations in certain key subgroups while avoiding, as much as possible, instabilities due to small cells.

The procedure used in 1978-79 has the following characteristics:

1. It produces a single ad-

justed weight for each individual.

2. It affords good control on the distribution of proportions of certain key variables.

3. It tends to produce stability of performance estimates.

4. It is relatively easy to implement.

Even though adjusted weights using this procedure differ slightly from the corresponding adjusted weights from the other procedures that have previously been employed, National Assessment intends to use weights obtained using the 1978-79 procedure for all future analyses of data assessed in earlier years. This is simply because we believe weights obtained through this procedure have the most desirable features.

The Current Smoothing Procedure

The first step in the 1978-79 smoothing procedure involved the partitioning of the population of age-class eligibles into the six smoothing cells given in Table E-1. The same cells were used for all ages.

Then, for each age and every year, the proportion of the population in each of the cells was estimated. For a given age and year, the proportion of the population in a particular cell was computed as the sum of weights of all respondents assessed in the given year who were of the specified age and who belonged in the cell, divided by the total of the weight of all respondents of the given age assessed in that year.

Each of the six cells was comprised of a sequence of estimated population proportions corresponding to the various years of assessment. Each such sequence of proportions was then smoothed by fitting robust/resistant lines. Using data from the U.S. Census Bureau and Current Population Survey, trends in enrollment by age and race and by age and region were obtained. The data from these surveys were ad-

---

TABLE E-1. Smoothing Cells Used for the 1978-79 Smoothing Procedure

Cell	Race	Region	Community Size
1	White	All	Big cities + fringe
2	White	All	Medium cities
3	White	All	Small places
4	Black	Southeast	All
5	Black	Not Southeast	All
6	Other	All	All

justed to correspond with National Assessment definitions as much as possible. The resistant lines within the smoothing cells were constrained to satisfy the trend from the U.S. Census and Current Population Survey data.

The final step in the smoothing procedure was to adjust the respondents' weights to be consistent with the smoothed proportions. Since each respondent takes only one booklet, the weight adjustments were done independently for each booklet. For a given age, year and booklet, population proportions using the original weights were obtained for each of the smoothing cells. Then the weights of all respondents of a given cell were multiplied by the ratio of the smoothed cell proportion to the proportion using the original weights. This produced the adjusted weights that were used in all analyses.

Adjustment of Weights by Users

The smoothed population proportions for 9-, 13- and 17-year-olds (in-school only) are given in Tables E-2, E-3 and E-4, respectively. The columns of each table represent the smoothing cells while the rows

---

TABLE E-2. Smoothed Frequencies From 10-Year Smooth by Smoothing Cell and Year for 9-Year-Olds

	Cell					
	1	2	3	4	5	6
Race	White	White	White	Black	Black	Other
Region	All	All	All	SE+	Not SE	All
Size of Community	BC+FR#	MC++	SP##	All	All	All
Year						
1970-71	0.3299	0.1203	0.3574	0.0557	0.0736	0.0631
1971-72	0.3232	0.1177	0.3647	0.0562	0.0743	0.0638
1972-73	0.3165	0.1152	0.3720	0.0568	0.0749	0.0646
1973-74	0.3098	0.1126	0.3793	0.0573	0.0756	0.0654
1974-75	0.3030	0.1101	0.3866	0.0579	0.0763	0.0661
1975-76	0.2963	0.1076	0.3938	0.0584	0.0770	0.0668
1976-77	0.2896	0.1050	0.4011	0.0590	0.0776	0.0676
1977-78	0.2829	0.1025	0.4084	0.0596	0.0783	0.0684
1978-79	0.2762	0.1000	0.4157	0.0601	0.0790	0.0691
1979-80	0.2694	0.0974	0.4230	0.0607	0.0797	0.0698

+SE = Southeast.
#BC+FR = big cities + fringes.
++MC = medium cities.
##SP = small places.

TABLE E-3. Smoothed Frequencies From 10-Year Smooth by Smoothing Cell and Year for 13-Year-Olds

	1	2	3	4	5	6
Race	White	White	White	Black	Black	Other
Region	All	All	All	SE+	Not SE	All
Size of Community	BC+FR#	MC++	SP##	All	All	All
Year						
1970-71	0.3327	0.1113	0.3748	0.0523	0.0679	0.0610
1971-72	0.3279	0.1106	0.3779	0.0524	0.0694	0.0618
1972-73	0.3232	0.1098	0.3180	0.0525	0.0709	0.0626
1973-74	0.3184	0.1091	0.3841	0.0526	0.0724	0.0634
1974-75	0.3137	0.1084	0.3872	0.0527	0.0739	0.0642
1975-76	0.3089	0.1076	0.3903	0.0528	0.0754	0.0650
1976-77	0.3042	0.1069	0.3933	0.0528	0.0770	0.0658
1977-78	0.2994	0.1062	0.3964	0.0529	0.0785	0.0666
1978-79	0.2946	0.1055	0.3995	0.0530	0.0800	0.0674
1979-80	0.2899	0.1047	0.4026	0.0531	0.0815	0.0682

+SE = Southeast.
#BC + FR = big cities + fringes.
++MC = medium cities.
##SP = small places.

TABLE E-4. Smoothed Frequencies From 10-Year Smooth by Smoothing Cell and Year for In-School 17-Year-Olds

	1	2	3	4	5	6
Race	White	White	White	Black	Black	Other
Region	All	All	All	SE+	Not SE	All
Size of Community	BC+FR#	MC++	SP##	All	All	All
Year						
1970-71	0.3634	0.1205	0.3670	0.0438	0.0581	0.0472
1971-72	0.3577	0.1199	0.3704	0.0444	0.0597	0.0478
1972-73	0.3519	0.1194	0.3738	0.0451	0.0614	0.0484
1973-74	0.3462	0.1188	0.3772	0.0457	0.0630	0.0491
1974-75	0.3404	0.1183	0.3806	0.0463	0.0647	0.0497
1975-76	0.3347	0.1177	0.3840	0.0470	0.0663	0.0503
1976-77	0.3290	0.1172	0.3874	0.0476	0.0679	0.0509
1977-78	0.3232	0.1166	0.3907	0.0482	0.0696	0.0515
1978-79	0.3175	0.1161	0.3941	0.0489	0.0712	0.0522
1979-80	0.3117	0.1155	0.3975	0.0495	0.0729	0.0528

+SE = Southeast.
#BC+FR = big cities + fringes.
++MC = medium cities.
##SP = small places.

represent the assessment year. For example, the smoothed population proportion of 9-year-olds in smoothing cell 2 (whites in medium cities) for 1972-73 is .1152.

To adjust respondent weights to be consistent with the smoothed proportions, the following procedures were followed:

1. For each booklet, respondents were classified according to smoothing cell, and the raw population proportions for each cell were obtained. For example, the raw proportions for a booklet given to 9-year-olds in smoothing cell 4 was the total of the weights of all 9-year-olds receiving the booklet who were black and in the Southeastern region, divided by the total of the weights of all respondents receiving the booklet.

2. For each booklet and smoothing cell, a weight adjustment factor as the ratio of the smoothed population proportion (for the appropriate age, year and smoothing cell) over the raw population proportion was obtained.

3. The adjusted weights for an individual were the product of that individual's original weight and the appropriate adjustment factor.

Changes in Smoothed Proportions as New Assessments Are Completed

Every time an assessment is completed, a new time point is added to each of the sequences of population proportions within the smoothing cells. This means that, even though robust/resistant procedures are used, the addition of a new point may somewhat change the values of smoothed proportions for prior years. Additionally, any changes in methodology will impact the estimates.

This means that the smoothed proportions, with the addition of the next assessment data, are apt to differ somewhat from the corresponding smoothed proportions without the new data. National Assessment has adopted the philosophy that the smoothed proportions, based on all currently available data using the best available algorithm, are the best available. Therefore, all subsequent analyses, for any year, will be done using this best available information, even though this may produce estimates that differ slightly from prior values.

GLOSSARY OF NATIONAL ASSESSMENT TERMS

Acceptable response. Any response to an exercise that demonstrates achievement of the objective measured by that exercise.

Administration time. The total time allowed on the paced audio tape for an exercise. (Includes the time allowed for the stimulus and the response.)

Administration timetable. Time periods during the school year when the various age groups are assessed. The time periods are:

October-December	13-year-olds
January-February	9-year-olds
March-May	17-year-olds

Age group or age level. Three age groups have been sampled in both art assessments: 9-year-olds, 13-year-olds and 17-year-olds attending school. Birth date ranges for each age group in each of the two assessments are as follows:

Assessment	Age 9	Age 13	Age 17
1974-75	1965	1961	10/57-9/58
1978-79	1969	1965	10/61-9/62

Assessment. The documentation of the progress in knowledge, skills and attitudes of American youth. Measures are taken at periodic intervals for each learning area, with the goal of determining trends and reporting the findings to the public and to the education community.

Assessment administrator. Individual employed to administer the assessment in participating schools.

Background questions. Questions about respondents' instructional experiences with art both in school and out of school were included in art item booklets. Standard background questions asked in every learning area are found on the back pages of the item booklets and include such things as level of parental education and reading materials in the home. Background questions used in the 1978-79 art assessment appear in Appendix B.

Booklet. Items (exercises) are presented to respondents in booklets. Booklets are designed to be scored by optical scanning machines. Each booklet contains: (1) instructions for answering items and sample items, (2) assessment items and (3) background questions. Each booklet contains approximately 30-35 minutes of assessment items and 10-15

minutes of introductory material and background questions.

Category (scoring). A classification of a response to an open-ended exercise. See scoring guide.

Category within a variable. A subclassification within a variable. For example, male and female are categories of the variable sex. See reporting groups.

Difficulty level. The percentage of acceptable responses to an exercise.

Exercise. A task designed to measure an objective. Because NAEP does not administer "tests," but instead describes educational achievement over time, the term "exercise" is often used instead of the term "item" or "test item." The terms "item" and "exercise" are used synonymously in this report.

Exercise booklet. See booklet.

Exercise part. See item part.

Exercise pool. The entire set of exercises prepared for a learning area. This set includes recycled exercises, exercises developed for previous assessments but not used due to exercise booklet or budgetary constraints and newly developed exercises.

Field test. A pretest of exercises to obtain information regarding clarity, difficulty levels, timing, feasibility and special administrative problems. Needed for revision and selection of exercises to be used in the assessment.

Grade in school. Results are reported for 9-year-olds in grades 3 and 4; 13-year-olds in grades 7 and 8; and 17-year-olds in grades 10, 11 and 12.

Group administration. Booklets were administered to groups of 10 to 25 students in 1978-79. In 1974-75, normal group size was 12 respondents. A paced audio tape was used to provide uniform instructions and oral presentation of exercises.

Hand scoring (scoring). The coding of responses in a format compatible with the optical scanning equipment being used. Multiple-choice exercises can be directly machine scored; however, responses to open-ended exercises must be coded in scoring ovals so that they can then be machine scored. See scoring guide.

ID number. An identification number referring to the unique number assigned to each respondent. This number is assigned to preserve the anonymity of each respondent. NAEP does not keep records of the names of any individuals.

Item. See exercise.

Item booklet. See booklet.

Item part. Each part of an item that asks a separate question. Parts may all pertain to one stimulus, such as a graph or a table, or may concern the same topic.

Jackknife. The name of the algorithm used by NAEP to estimate standard errors of percentages and other statistics.

Learning area. One of the areas assessed by National Assessment: reading/literature, writing, mathematics, science, citizenship/social studies, art, music, career development. Also called "subject area."

Level of parental education. These levels are described in Appendix A.

Modal grade. The grade in which the majority of each in-school age group is enrolled. For 9-year-olds, the modal grade is grade 4; for 13-year-olds, grade 8; and for 17-year-olds, grade 11.

Objective. A desirable education goal agreed upon by scholars in the field, educators and concerned lay persons, and established through the consensus approach.

Objectives redevelopment. After the initial assessment of a learning area, one of the first steps in preparing for reassessment is a review of the learning-area objectives. This is carried out by scholars in the field, educators and concerned lay persons. These reviews may result in revision, modification or total rewriting of the learning-area objectives to reflect current curricular goals and emphases; they may also result in the endorsement of the objectives from the previous assessment as adequate for the next assessment.

Open-ended exercise. A nonmultiple-choice exercise that requires some type of written or oral response.

Paced audio tape. A tape recording that accompanies each booklet to assure uniformity in administration. Instructions and exercises are played back from the tape recording so that reading difficulties will not interfere with an individual's ability to respond. Response time is included on the tape.

Primary sampling unit (PSU). First-stage sampling units, typically a county or a group of contiguous counties.

Principal's questionnaire. A data collection form given to school officials. The officials respond to questions concerning enrollments, size of the community, occupational composition of the community, and so forth. Samples of these questionnaires are found in Appendix B. See also supplementary principal's

questionnaire.

PSU. See primary sampling unit.

Public-use data tapes. Computer tapes containing respondent-level exercise and background/demographic data and machine-readable documentation. These tapes are available for use by external researchers wishing to do secondary analyses of National Assessment data.

Racial/ethnic category. For the art assessments, results are reported for whites, blacks and Hispanos. (Because of a small sample size, only average results for Hispanos can be reported by National Assessment.)

Receipt control. Procedures implemented by scoring staff to check in and screen materials from the field. Information gained from receipt control procedures is relayed to the administrative staff so that any errors may be corrected.

Recycled exercises. The set of exercises that is kept secure from one assessment to the next that will be used to measure changes (growth, stability or decline) in performance for the learning area.

Region. One of four geographical regions used in gathering and reporting data: Northeast, Southeast, Central and West. States included in each region are shown in Appendix A.

Released exercise. An exercise for which results and exercise text have been reported to the public.

Released exercise set. A set of released exercises, including documentation and scoring guides, that can be purchased from National Assessment.

Reporting groups. Categories of variables for which National Assessment data are reported. Variable categories are defined in Appendix A.

Rescore. If an open-ended exercise was scored under different conditions than presently held or if passage of time might affect scoring, responses from a previous assessment may be rescored at the same time that responses from a later assessment are scored. Responses from an earlier assessment also may be held and not scored so that they can be scored with responses from a later assessment.

Respondent. A person who responds to the exercises in an assessment booklet.

Response options. Different alternatives to a multiple-choice question that can be selected by the respondent.

Review conference. A conference held to review the objectives of a learning

area to ensure their acceptance as measures of the objectives by scholars, educators and lay persons or to review exercises for racial, ethnic, social or regional bias.

Sample. National Assessment does not assess an entire age population but rather selects a representative sample from the age group to answer assessment items. (See Chapter 4 for a description of National Assessment sampling procedures.)

Scoring guide. A guide for hand scoring an open-ended exercise that specifies descriptive or diagnostic categories by giving definitions and sample responses.

Scoring ovals. Scannable ovals printed beside multiple-choice options and printed at the bottom of the page for open-ended exercises (to be used in hand scoring). When ovals are marked, they can be scored by machine and responses then can be recorded by computer.

Sex. Results are reported for males and females.

Size of community. Results are reported for four size-of-community categories: big cities, fringes around big cities, medium cities and small places. These categories are defined in Appendix A.

SMSA. Standard Metropolitan Statistical Area. SMSAs are economic and social units defined by the U.S. Bureau of the Census.

Standard error. A measure of sampling variability for a statistic. Because of NAEP's complex sample design, standard errors are estimated by jackknifing first-stage sample estimates.

Stem. The portion of an exercise that states the problem or asks the question.

Stimulus. For art exercises, this will be an aural and/or a visual stimulus used as part of the stem.

Subject area. See learning area.

Subpopulation or subgroup. Groups within the national population, such as males and females, for which data are reported.

Supplementary principal's questionnaire. A data collection form given to school officials. On this form, officials are asked to respond to questions concerning course offerings, materials and staffing specific to the learning area being assessed. A sample of this questionnaire is found in Appendix B. See also principal's questionnaire.

Tapescript. A script prepared for the announcer to use in

producing the paced tape. It indicates exactly what is to be read or not read aloud to the students and indicates the amount of response time allowed for each exercise. See paced audio tape.

Timing. Most NAEP exercises are administered with a paced audio tape to standardize data collection conditions. The tape includes the amount of time students are allowed to respond to each exercise.

Type of community. Results are reported for three type-of-community categories: disadvantaged urban, advantaged urban and rural. Definitions of these categories are found in Appendix A.

User tape. See public-use data tape.

Variable. A classification of respondents. Standard reporting variables are: region, sex, race/ethnicity, level of parental education, size of community, type of community and grade in school.

Weight. A multiplicative factor equal to the reciprocal of the probability of a respondent being selected for assessment with adjustment for nonresponse -- an estimate of the number of persons in the population represented by a respondent in the sample. Theoretically, the sum of the weights for all respondents at an age level is equal to the number of persons in the country at that age level.

BIBLIOGRAPHY

Art Objectives, no. 06-A-10, 1974-75 Assessment. Denver, Colo.: National Assessment of Educational Progress, Education Commission of the States, 1971. ERIC no. ED 051 255. ISBN 0-89398-001-3.

Art Technical Report: Exercise Volume, Report no. 06-A-20, 1974-75 Assessment. Denver, Colo.: National Assessment of Educational Progress, Education Commission of the States, 1978. ISBN 0-89398-006-4.

Art Technical Report: Summary Volume, Report no. 06-A-21, 1974-75 Assessment. Denver, Colo.: National Assessment of Educational Progress, Education Commission of the States, 1978. ERIC no. ED 155 125. ISBN 0-89398-007-2.

Attitudes Toward Art, Report no. 06-A-03, 1974-75 Assessment. Denver, Colo.: National Assessment of Educational Progress, Education Commission of the States, 1978. ERIC no. ED 166 122. ISBN 0-89398-005-6.

Chromy, J. and D. Horvitz. "Structure of Sampling and Weighting," 1969-70 Science: National Results and Illustrations of Group Comparisons, Report no. 1, 1969-70 Assessment. Denver, Colo.: National Assessment of Educational Progress, Education Commission of the States, 1970. ERIC no. ED 055 786. ISBN 0-89398-276-8.

Design and Drawing Skills, Report no. 06-A-01, 1974-75 Assessment. Denver, Colo.: National Assessment of Educational Progress, Education Commission of the States, 1977. ERIC no. ED 141 249. ISBN 0-89398-002-1.

Final Report: In-School Field Operations and Data Collection Activities, National Assessment of Educational Progress, For the Period of Oct. 1, 1974 Through Sept. 30, 1975. Research Triangle Park, N.C.: Research Triangle Institute, 1975.

Final Report on National Assessment of Educational Progress Sampling and Weighting Activities for Assessment Year 10. Research Triangle Park, N.C.: Research Triangle Institute, 1980.

Final Report on National Assessment of Educational Progress Sampling and Weighting Activities for Year 06. Research Triangle Park, N.C.: Research Triangle Institute, 1976.

Final Report on Year 10 In-School Field Operations and Data Collection Activities, National Assessment of Educational Progress. Research Triangle Park, N.C.: Research Triangle Institute, 1979.

Folsom, Ralph E. National Assessment Approach to Sampling Error Estimation, Sampling Error Monograph 25U-796-5. Research Triangle Park, N.C.: Research Triangle Institute, 1977.

Kalsbeek, W.D. et al. No Show Analysis, final report. Raleigh, N.C.: Research Triangle Institute, 1975.

Knowledge About Art, Report no. 06-A-02, 1974-75 Assessment. Denver, Colo.: National Assessment of Educational Progress, Education Commission of the States, 1978. ERIC no. ED 151 270. ISBN 0-89398-004-8.

Miller, R.G. Jr. "A Trustworthy Jackknife," Annals of Mathematical Statistics, no. 35, 1964.

Miller, R.G. Jr. "Jackknifing Variances," Annals of Mathematical Statistics, no. 39, 1968.

Miller, R.G. Jr. "The Jackknife -- A Review," Biometrika, no. 61, 1974.

Mosteller, F. and J.W. Tukey. Data Analysis and Regression. Reading, Mass.: Addison-Wesley, 1977.

Mosteller, F. and J.W. Tukey. "Data Analysis Including Statistics," in Handbook of Social Psychology, 2nd ed., edited by E. Aronson and G. Lindzey. Reading, Mass.: Addison-Wesley, 1968.

Rogers, W.T. et al. "Assessment of Nonresponse Bias in Sample Surveys: An Example From National Assessment," Journal of Educational Measurement, vol. 14, no. 4, 1977.

Supplement to 1971 Art Objectives, no. 10-A-11, 1978-79 Assessment. Denver, Colo.: National Assessment of Educational Progress, Education Commission of the States, 1971. ISBN 0-89398-010-2.

The Second Assessment of Art, 1978-79: Released Exercise Set, no. 10-A-25. Denver, Colo.: National Assessment of Educational Progress, Education Commission of the States, 1980. ERIC no. ED 186 331. ISBN 0-89398-011-0.

Tukey, John W. Exploratory Data Analysis. Reading, Mass.: Addison-Wesley, 1977.

Tukey, John W. "Some Considerations on Locators Apt for Some Squeezed-Tail (or Stretched-Tail) Parents" (paper produced in connection with research at Princeton University, supported by the Army Research Office), summer 1975.

☆ U. S. GOVERNMENT PRINTING OFFICE 1981 - 576-063/78 Reg. 8

NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS
Education Commission of the States

Robert D. Ray, Governor of Iowa, Chairperson,
Education Commission of the States
Robert C. Andringa, Executive Director,
Education Commission of the States
Roy H. Forbes, Director, National Assessment

All National Assessment reports and publications are available through NAEP offices at the address shown at the bottom. Reports ordered from NAEP are processed within three days.

Some of the more recent reports are also available from the Superintendent of Documents, and most all are available from ERIC. Write to the National Assessment office for a current publications list.

National Assessment reports related to this report:

MUSIC
1st Assessment (1971-72)

03-MU-01	The First National Assessment of Musical Performance, February 1974	$ 1.00
03-MU-02	A Perspective on the First Music Assessment, April 1974	1.00
03-MU-03	An Assessment of Attitudes Toward Music, September 1974	1.10
03-MU-00	The First Music Assessment: An Overview, August 1974	1.00
03-MU-20	Music Technical Report: Exercise Volume, December 1975	25.00
03-MU-21	Music Technical Report: Summary Volume, November 1975	4.40

NOTE: A cassette supplementing the music reports including musical stimuli and actual performance by 9-, 13- and 17-year olds is available for $1 [illegible].

2nd Assessment (1978-79)

10-MU-01	Music 1971-79: Results From the Second National Music Assessment, November 1981	7.00
10-MU-25	The Second Assessment of Music, 1978-79, Released Exercise Set, April 1980	15.20
10-MU-40	Procedural Handbook: 1978-79 Music Assessment, December 1981	7.70

NOTE: A cassette tape with music stimuli for released exercises supplementing Report no. 10-MU-25 is now available for $2.00.

(Continued on Inside Back Cover)

(Continued From Inside Front Cover)

ART

1st Assessment (1974-75)

06-A-01	Design and Drawing Skills, June 1977	3.70
06-A-02	Knowledge About Art, January 1978	2.15
06-A-03	Attitudes Toward Art, May 1978	2.65
06-A-20	Art Technical Report: Exercise Volume, January 1978	27.50
06-A-21	Art Technical Report: Summary Volume, June 1978	6.80

2nd Assessment (1978-79)

10-A-01	Art and Young Americans, 1974-79: Results From the Second National Art Assessment , December 1981	8.90
10-A-25	The Second Assessment of Art, 1978-79, Released Exercise Set, April 1980	13.65
10-A-40	Procedural Handbook: 1978-79 Art Assessment, December 1981	8.60

BACKGROUND REPORT

03/04-GIY	General Information Yearbook. A condensed description of the Assessment's methodology, December 1974	2.75

In addition to the above reports, National Assessment has produced reports in the areas of social studies, citizenship, writing, literature, reading, mathematics, science and career and occupational development. A complete publications list and ordering information are available from the address below.

NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS
Suite 700, 1860 Lincoln Street
Denver, Colorado 80295